UNIVERSITY OF PADOVA

Department of General Psychology

Master Degree in Cognitive Neuroscience and Clinical Neuropsychology

Final dissertation

A Machine-Learning-Based Investigation of Schizophrenia Using Structural MRI

Supervisor
Professor Hugo Schnack*

Co-supervisor
Professor Cristina Scarpazza

Candidate: Ayşe Dilara Türkmen
Student ID number: 1233878

Academic Year 2021-2022

* University Medical Center Utrecht, Department of Psychiatry, Utrecht University

Table of Contents

Abstract

Schizophrenia is a serious mental health concerns that affects 1% of the population (Jones et al., 2005). This study aimed to create objective tools that can correctly classify people with schizophrenia according to their diagnosis, predominant symptoms, illness duration, and illness severity based on their structural brain imaging variables. 1087 brain images (700=healthy controls, 387=people with schizophrenia) included in the analysis. Support Vector Machines, random forests, logistic regression, and XGBoost were used for diagnostic classification and reached 71% of maximum accuracy. Sulcal width was found to be the most important brain imaging variable that differed between groups. Support vector machines and random forests were used to classify patients according to their predominant symptoms and these classifications reached a maximum accuracy of 66%. Support vector machines could correctly classify people with schizophrenia according to their illness duration with a 75% accuracy and according to their illness severity with 69%. The result of the study shows that using machine learning methods, it is possible to create objective tools for schizophrenia that can be later used in clinics.

Keywords: Schizophrenia, Structural MRI, Machine Learning Classification

List of Abbreviations

APA= American Psychiatric Association

AUC=Area Under Curve

CT= Cortical Thickness

DSM=Diagnostic and Statistical Manual

GM= Grey Matter

HC group= Healthy control group

IC= Intracranial Volume

ICD= International Classification of Diseases

LGI=Local Gyrification Index

LR= Logistic Regression

MRI=Magnetic Resonance Imaging

RF= Random Forest

ROC= Receiver Operating Characteristic

SI= Sulcation Index

STG= Superior Temporal Gyrus

SVM= Support Vector Machines

SZ group= Schizophrenia group

TNR=True Negative Rates

TPR= True Positive Rates

WM=White Matter

XGBoost= Extreme Gradient Boosting

Acknowledgements

I would like to thank my external supervisor Hugo Schnack for his guidance through his extensive knowledge on the topic and valuable feedbacks. I am grateful to Joost Janssen for preparing the dataset I used in my study. Thanks to my internal supervisor Cristina Scarpazza for her support.

Special thanks to my parents Gülhan and Mehmet Türkmen, my sister and my friends for showing emotional support at every stage of my master's degree even though we are 1800 km apart.

CHAPTER 1

INTRODUCTION

Schizophrenia is a serious disorder with a heterogeneous clinical manifestation characterized by psychological, behavioral and cognitive symptoms (Hopkins et al., 2018). Around 1% of the population is affected by the disorder (Jones et al., 2005) and this amount is expected to rise due to population growth and ageing (Charlson et al., 2018). Prevalence of schizophrenia was reported to range from 4 to 7 per 1000 people in a large systematic review study (Saha et al., 2005). Median incidence rate for the disorder is 15.2 per 100000 (McGrath et al., 2004).

Schizophrenia is considered a disabling disorder due to its early onset, chronic course, and related cognitive, physical, and psychological impairments (Mueser & Jeste, 2008). Long years of intervention efforts did not change the fact that schizophrenia is one of the leading reasons of disability, in fact, Samuel et al. (2018) reported only 2% of the people with schizophrenia were completely independent in their daily lives, while the rest had problems with numerous daily life activities from shopping to preparing meals. Another study showing dramatic effects of the disorder was reported by Salomon and colleagues (2012), where they calculated disability weights for 220 disorders on a scale from 0 to 1 in a way that 0 represents no loss of health and 1 represents a health loss equivalent to death. While a mild diarrhoea's disability weight calculated to be 0.061, a terminal phase cancer's disability weight calculated to be 0.508. Out of those 220 disorders, the highest disability weight belonged to acute state schizophrenia with 0.756.

Moreover, rates of mortality are alarmingly high for people with schizophrenia compared to general population with a reduction of 10-25 years of life expectancy and possible reasons behind this can be stated as high suicide rates among patients, unhealthy lifestyle, negative

consequences of antipsychotics, and inadequate treatment for comorbid physical illnesses in patients (Laursen et al., 2012). Comorbid physical illnesses account for %60 of the premature deaths that are not related to suicide in schizophrenia patients (Lambert et al., 2003). One possible reason was stated by Briskman et al. (2012) as the results of their study showed that comorbid physical illnesses were less likely to be reported in psychiatric patients compared to the control group, psychiatric group were less likely to receive medication for those problems and they were less likely to be educated about lifestyle and medical interventions related to their comorbid physical illnesses.

As could be expected, schizophrenia is related to high economic burden, Chong et al. (2016) explained yearly costs of schizophrenia differed from US$94 million to US$102 billion. The economic burden of the disease differed from 0.02% to 1.64% of the gross domestic product. Authors argues that this extreme cost suggests unsatisfactory health care for patients.

1.1 Risk Factors of Schizophrenia

Schizophrenia is a highly heritable disorder (Ripke et al., 2014). However, literature suggests that there is an undeniable effect of non-genetic factors as well (Stilo & Murray, 2019). In this section, non-genetic risk factors of schizophrenia will be discussed.

Many studies showed the problems during fetal development is related to higher risk of schizophrenia (Lipner et al., 2019). In their study, Pugliese et al. (2019), investigated the relationship between maternal distress, medical illnesses, obstetric complications, and some mental disorders including schizophrenia. Increased risk of schizophrenia in later life was associated with psychological stress, inadequate weight gain, and infections during pregnancy.

Socioeconomic status investigated comprehensively as a risk factor for schizophrenia. Low individual and community level socioeconomic status at birth is a risk factor for

schizophrenia (Werner et al., 2007). One meta-analysis study found that prevalence of schizophrenia was 10% among homeless people (Ayano et al., 2019). Interestingly, the results of one study done in China showed that the prevalence of schizophrenia was highest for people with low individual-level socioeconomic status living in high community-level socioeconomic status (Luo et al., 2019). Immigration is stated to be another risk factor for schizophrenia, results of meta-analysis show that not only first-generation immigrants, but also second-generation immigrants have an increased risk of schizophrenia (Bourque et al., 2011). Experienced discrimination and social exclusion are seen to be the core causes of this increased risk (Henssler et al., 2020). Furthermore, low socioeconomic status of immigrants has shown to be a factor as the difference of the risk for schizophrenia between minorities and majority group decreases when socio-economic indicators are adjusted (Hjern et al., 2004). Childhood trauma and social adversities, isolation, living in an urban environment, and cannabis and other substance use are the other stated risk factors of schizophrenia (Stilo & Murray, 2019).

1.2 Symptoms of Schizophrenia

Records of schizophrenia-like disorder dates to Vedas of ancient Hindus (1400 BC), which document a condition that is brought by devils, which causes effected person to be nude, filthy, confused and lacks self-control (Howells 1991; Zilborg and Henry 1941; Lewis, 1966; Gottesman 1991; as cited in Adityanjee et al., 1999). The first elaborated description of the disorder dates to 18th century and later in the 19th century, it was defined as an early dementia by Benedict Augustine Morel, who used the term “precocious dementia” (Mueser & Jeste, 2008). Modern conceptualization began with Emil Kraepelin who integrated many conditions into a single disorder with early onset, poor prognosis, and “psychic” and “bodily symptoms (Mueser & Jeste, 2008). In his book “Dementia Praecox oder Gruppe der Schizophrenien”, Bleuler

(1911), used the word schizophrenia for the first time. The word comes from Greek words meaning “split” and “mind” (Moskowitz, 2008). Later, extensive research has been made to profoundly understand the symptoms of schizophrenia for a better diagnosis and understanding of the disease. Current diagnostic criteria and understanding of symptomatology of the disease will be discussed in the following section.

According to Diagnostic and Statistical Manual of Mental Disorders 5 (5th ed.; DSM–5; American Psychiatric Association, 2013), a person should be diagnosed with schizophrenia if they show at least two of the following: delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior and negative symptoms for a significant portion of time during 1-month period. At least one of the two of them should be delusions, hallucinations, or disorganized speech (APA, 2013).

1.2.1 Delusions

John Nash, a respected mathematician from West Virginia (Weiden, 2002), believed that invisible powers from space, or maybe foreign governments, were trying to communicate him using New York Times by encrypting messages. Only target was him and he needed serious analysis to find the messages, he was allowed to learn the secrets of the world (Nasar, 1998). In DSM-5, delusions are defined as rigid beliefs that the person who experiences them are not open to change them even in the light of contradictory proof. Delusions are divided into as bizarre and non-bizarre delusions, the former refers to the delusions that have no basis in reality, whereas latter refers to the delusions that can happen such as a cheating spouse. The content of the delusions may vary (APA, 2013). Like many other people with schizophrenia, John Nash was suffering from bizarre delusions. Delusions of clients have been hard to understand since the clinicians are generally facing rational people holding an impossible belief. (Feyaerts et al.,

2021). Research been done to understand how these beliefs are formed. Langdon and Coltheart (2000) stated there should be two different types of problems in cognition to result in delusions. First of them is a damage to sensory and/or attentional mechanism which results in abnormal perception. This damage would explain the bizarre content, however, it is not enough on its own as an explanation. This leads us to their second proposed problem in the normal cognition which is the failure of evaluation of normal belief.

To illustrate better, an example can be experience of passivity or delusions of control. Patients with such experience believe their actions are controlled by an external agent. (Schnell et al., 2008). In everyday life, it is important for us to know the difference between movements that is willingly initiated by us and the movements due to external effects. Such differentiation exists in the action system as discussed by Frith (1987), there are two routes, and the first route involves perceiving a stimulus and an action is made as it is required by the specific stimulus. The second route, on the other hand, starts when the subject has a goal, and an action is created to attain that particular goal. Frith suggested existence of a monitoring mechanism for these routes, which detects if the action is initiated by a stimulus or a goal. The failure of this mechanisms results in delusion of control in patients since the distinction between the actions started by the stimuli or goal is ambiguous. Some researchers did experiments to understand the mechanisms more extensively. In their study, Schnell et al. (2008) selected participants with and without passivity symptoms. They were asked to perform a visuomotor task where they had to control the movements of a car, but it was disrupted by the computer at random times. Participants were instructed to stop controlling as soon as they realize computer was in charge for the movement. The results showed that there was a correlation between passivity symptoms and impaired behavioral monitoring performance. Going back to the theory by Langdon and

Coltheart (2000), this problematic monitoring can be seen as the first factor of delusion formation. However, this disruption is not sufficient to explain delusion formation. If perceptual aberration was sufficient for delusion formation, every person who suffers from that would experience delusions. Visiting the phenomena called phantom limb experience, it is not the case. Patients who had a limb amputation may have a vivid experience that the limb is still there, and it is painful (Langdon & Coltheart, 2000). This sensation thought to be a result of reorganization in the somatosensory cortex (Makin & Flor, 2020). However, these patients do not have the delusion of their limb being there. Thus, only the first factor is not enough to explain delusion formation (Langdon & Coltheart, 2000). An impaired belief evaluation should be added to the equation (Coltheart et al., 2011). However, the exact nature of this factor is not specified and further research should be done for this aim (Feyaerts et al., 2021).

1.2.2 Hallucinations

Hallucinations are defined as perception-like experiences that happen in the absence of an external stimuli. (APA, 2013). Although sensory modality for hallucinations may differ, auditory hallucinations are far more common than any other modality, which followed by visual and cenesthetic hallucinations. Tactile, olfactory, and gustatory hallucinations are less common (Bauer et al., 2011).

A common mechanism underlying hallucinations was suggested by Behrendt and Young (2004) as antipsychotic drugs are successful at controlling hallucinations in different sensory modalities, and this is not an exception for schizophrenia patients. Many scientists with different approaches tried the clarify the etiological factors behind hallucinations (for detailed review, see Chaudhury et al., 2009). Similar to stated monitoring problems for delusion formation, many researchers found that schizophrenia patients who suffer from verbal and auditory hallucinations

were prone to mistakenly attribute their own voice to an outside source (Bentall et al., 1991; Johns et al., 2001, 2006; Woodward et al., 2007; Costafreda et al., 2008; Brébion et al., 2016).

Another proposal is that the deficits in the efference in the copy system which is when there is an upcoming action, an efferent copy warns the sensory areas about this upcoming action and labels the action as self-generated. A deficit in this system may result in externalizing bias and that makes patients more prone to attribute a stimulus to an outside source (Lieberman et al, 2020).

Another approach is predictive coding framework which integrates both perceptual and cognitive accounts for hallucinations (McCleery et al., 2018). Predictive coding accounts the brain as a hierarchy with a goal of maximizing proof for its model of the world. That happens by constantly comparing prior beliefs to sensory data. The mismatch is predictive errors and that leads to updating the model (Sterzer et al., 2018). Another aid for minimizing the errors is active inference which refers to actively selecting sensory proof to minimize errors (Friston, 2010). Fletcher and Frith (2009) suggested problematic predictive coding may lead to hallucinations and delusions in patients with schizophrenia. This happens when false prediction errors propagate, the model cannot update itself, resulting in hallucinations, and bizarre beliefs.

1.2.3 Disorganized Speech

Disorganized speech, or formal thought disorder (Kerns & Berenbaum, 2002), is defined as speech that is hard to understand with flawed organization (Becker et al., 2012). It is one of the central signs of schizophrenia (Kerns & Berenbaum, 2002) and as discussed by Andreasen (1979), its clinical manifestation can include switching topics unusually, using made up words (handshoes for the word gloves), and being incomprehensible (as cited in Kerns & Berenbaum, 2002).

Semantic memory impairments are related to the formal thought disorder in patients with schizophrenia (eg. Assaf et al., 2006; Jamadar et al., 2013; Stirling et al., 2006). The impairments of patients, proposed to be caused by overactivation and/or disorganization of semantic memory (Tan et al., 2015). Overactivation hypothesis states that, loose associations and disrupted thinking is the results of faster and more spread initial activation in the semantic networks (Niznikiewicz, 2008). Since this initiation propagates too far, the person leaves with too many word choices which, in turn, results in the problematic speech (Niznikiewicz et al., 2010). Disorganization hypothesis of semantic memory states that the core problem in the semantic memory within patients is the compromised semantic store but there is mixed evidence for both approaches (Tan et al., 2015). Furthermore, a body of research suggest that formal thought disorder in schizophrenia is more related to executive deficits than semantic deficits (Barrera et al., 2005). There is also research that favors both (e.g. Stirling et al., 2006). Further research is needed to fully understand the cognitive mechanisms behind disorganized speech.

1.2.4 Grossly disorganized and catatonic behavior

In DSM, these symptoms suggested to be seen in different ways from childlike silliness to unpredictable agitation. Catatonic behavior is defined as diminished reactivity to the environment. This can manifest itself as a rigid and bizarre posture, resistance to instructions or lack of verbal and motor responses (APA, 2013)

Although neurological background behind the abnormal motor behavior is highly tested, only a few studies were done to investigate cognitive correlates of these behaviors. With that aim Dean (2020) did an experiment with people who have schizophrenia with and without catatonic symptoms. Participants' cognitive abilities such as verbal fluency and processing speed are tested. Increased difficulties were found for patients with catatonic symptoms compared to

patients without catatonic symptoms. The difference was significant even after controlling the medication. No structural brain differences were detected between patients with and without catatonia. Thus, they suggested that cognitive domains can be a useful tool for differentiation among patients.

1.2.5 Negative Symptoms

Before talking about negative symptoms, it is important to state positive/negative symptom dichotomy that is commonly used for schizophrenia. Crow (1980) labelled positive symptoms as presence of abnormal psychological features such as hallucinations, delusions and thought disorder, whereas negative symptoms are defined as the absence of cognitive, and psychological functions or attributes that are generally present (Lieberman et al., 2020). Two main negative symptoms that are common in people with schizophrenia are diminished emotional expressions and avolition (APA, 2013). In their review article, Marder and Galderisi (2017), proposed five constructs which should be considered as negative symptoms. These are blunted affect, avolition, anhedonia, alogia, and asociality. Blunted affect (diminished emotional expressions) can be detected in a patient from the reduction in expressions of emotions in the face, movement of the head and hands, eye contact and intonation of speech (APA, 2013). This reduction in emotional expressions could not be attributed to diminished emotional experience of patients with schizophrenia (Kring & Moran, 2008). Decreased interest in initiation and perseverance of goal-directed activity is called avolition (Marder, & Galderisi, 2017). Reduction in speech is called alogia and reduction in the ability of experiencing pleasure is called anhedonia (APA, 2013). Lastly, as implied, asociality is defined as reduced interest in social activity (Marder, & Galderisi, 2017).

There is also a distinction made between negative symptoms as primary and secondary negative symptoms. Primary negative symptoms are considered as core features of the disease whereas secondary negative symptoms result from factors other than schizophrenia, or they are the result of particular symptoms of schizophrenia (Carpenter et al., 1985). This distinction is very important for the development of treatment strategies. For instance, the reason of asociality could be that the patient is uninterested, but it can be also a because of depression or anxiety (Kirkpatrick, 2014).

Around 25% of patients with schizophrenia spectrum disorder experience primary negative symptoms that are persistent over time (Chang et al., 2011). Rector et al. (2005), summarized cognitive processes behind negative symptoms in schizophrenia patients. They stated one of them is low expectancy for pleasure, when given the opportunity for pleasurable activities, a frequent answer from patients was “What’s the point?” or “It’s too much work”. They explained this by their low expectancy for pleasure in general. In one study, it has been showed that patients were asked how much time they spend on social, self-care and recreational activities and they were asked how much pleasure they expect to get from those activities. Patients reported lower levels and it mediated engaging in those activities (Pillny et al., 2020). Interestingly, even though patients underestimate the pleasure they would get by engaging in those activities, they show little impairment in pleasure when they actually engage in these activities (Rector et al., 2005). Cognitive model of negative symptoms is also supported by the finding of limited belief of success in patients. Patients’ beliefs about how likely they would be successful was related to reduced negative symptoms (Luther et al., 2015). Even when they accomplish their goals, they tend to consider their success to be not sufficient and this expectation of failure leads to less motivation for goal-directed behavior (Rector et al., 2005).

Herbener et al. (2008) showed that, during viewing positive images, patient group did not differ from controls, however, 24 hours later they could not retrieve those experiences. Thus, decreased motivation might be explained by problematic retrieving, updating, and maintaining the mental representation of value (Strauss et al., 2014). Furthermore, neuropsychological evaluations shows that schizophrenia patients with negative symptoms shows deficit in executive function and psychomotor speed, even though a global cognitive impairment difference is not detected (Bryson et al., 2001).

1.3 Neuroscientific Findings of Schizophrenia

Extensive research has been done to understand why patients with schizophrenia show such symptoms and what is the relationship between this disorder and the brain. In this section brain abnormalities found in patients with schizophrenia will be discussed. Main focus will be the structural brain abnormalities in schizophrenia since it is also the main focus of the current study. Later, functional, and neurochemical abnormalities will be briefly discussed.

1.3.1 Structural Brain Abnormalities in Schizophrenia

Many structural brain abnormalities found at every stage of schizophrenia (Zhao et al., 2018). Wright and colleagues (2000) did a meta-analysis including 58 studies on the structural abnormalities of schizophrenia. Whole brain volume of patients was lower compared to healthy control group. Patients' whole brain volume was founded to be 98% of a healthy person's brain volume (100%), similarly to whole brain, grey and white matter volumes were lower in patients. Relative grey matter volume of patients was founded to be 98% and relative white matter volume of patients was 99%. The biggest difference was observed in ventricular volume differences. The relative total ventricular volume of patients was founded to be 126%. The same trend showed

itself for every ventricular subdivision. The biggest volume difference was the volume of left lateral ventricle, which was 130% in patients.

Since lateral ventricular enlargement is one of the most stated findings in schizophrenia patients, a meta-analysis on this enlargement was made by Kempton et al. (2010). They included 13 longitudinal brain imaging studies that investigated lateral ventricular volume changes. A continuous increase in lateral ventricular volume was seen in schizophrenia patients, which was 3 to 4 times more than lateral ventricular volume changes due to healthy aging. Lateral ventricular enlargement can be even seen in patients with 20 years illness duration (Hulshoff Pol & Kahn, 2008). Bigger ventricle volume increase is associated with longer duration of psychosis (Cahn et al., 2009). Third ventricle but not lateral ventricle volume was associated with deficits in many cognitive abilities including attention and frontal functioning (Bornstein et al., 1992). Underlying pathophysiological processes behind ventricular enlargement is still not clear, however, ventricle enlargement founded to be a common finding in diseases that are neurodegenerative, thus this fact supports neurodegenerative hypothesis of schizophrenia (Svancer & Spaniel, 2021). Furthermore, increased ventricular enlargement in patients who do not use antipsychotics and first-episode patients may support the hypothesis of dopaminergic hyperactivity (Meduri et al., 2010).

As stated, grey matter differences between patients and healthy controls were more prominent than white matter changes (Wright et al., 2000). A meta-analysis of longitudinal MRI studies showed that, over time, patients show a greater reduction in cortical grey matter volume, however, this reduction grey matter volume is not even, some areas are affected more than the others (Vita et al., 2012). Reduction in grey matter volume of superior and medial temporal gyri,

anterior cingulate, thalamus, frontal lobe, hippocampus and amygdala are suggested (Glahn et al., 2008; Honea et al., 2005; Ohi et al., 2016).

Volume differences of superior temporal gyrus (STG) and its subregions repeatedly found in patients with schizophrenia (Sun et al., 2009). STG is an important region in temporal lobe, which includes primary and association auditory cortex, and network of connection to temporal limbic regions in the brain which has an important part in interpretation, production, and self-monitoring of language. (Kim et al., 2003; Sun et al., 2009). In their study, Rajarethinam et al. (2000) found that patients had reduced volume of left anterior STG, which had a negative relationship with the experienced severity of hallucinations among patients. They did not find any significant difference between patients and controls for left posterior STG, however, it also had a negative relationship with the experienced severity of thought disorder among patients. There was also an asymmetry in patients at left and right anterior STG, which was not the case for healthy controls. These results support the proposed primary auditory dysfunction in the anterior and middle STG, and auditory association cortex dysfunction in posterior STG, which may lead to abnormalities in auditory perception and organization of thought. McKinney et al. (2017), in their study investigating post-mortem brains of people with schizophrenia, found that there is a relationship between DNA methylation and dendritic spine density in STG and this relationship is disrupted in patients. They also suggested BAIAP2 and DLG1 genes to be candidate for mediating this disrupted relationship.

A significant reduction of anterior cingulate volume has found in patients with schizophrenia (Baiano et al., 2007). Abnormalities in anterior cingulate cortex (ACC) in patients might be a neurobiological basis of disease manifestation of schizophrenia since the region is involved in many important cognitive and emotional functional processes that leads to goal-

directed behavior (Fornito et al., 2009). Fujiwara and colleagues (2007) designed an experiment to investigate the relationship between structural abnormalities in the brain, their psychopathology and social cognition. Their task tested the ability of emotion attribution to facial expression and story characters of patients. They found a relationship between right ACC and positive symptoms severity of patients and another relationship was found between left paracingulate sulcus and negative symptoms severity of patients. An additional significant relationship was found between both left and right ACC volumes and the performance of two tasks. First task involved matching emotional facial expressions with emotional verbal labels and for the second one, participants needed to match appropriate emotional expression for given stories. These results show that pathology of ACC correlates with psychopathology and social cognition problems of schizophrenia. Due to the importance of anterior cingulate cortex on empathy, another study is done to investigate correlation between ACC volume and empathy processes. Their results showed that pathology of some ACC subdivisions have an impact on empathic disabilities in female schizophrenia patients (Fujiwara et al., 2008).

A meta-analysis by Adriano et al. (2010), found a significant bilateral thalamus volume reduction in both chronic and first-episode patients. Thalamus is a very important region in the brain, it has connections with most parts of the brain due to its location. Moreover, it is a sensory relay station, thus very vital for perception, but also its functions effect many aspects in cognition including memory, attention, and consciousness (Ward, 2013). One study showed that reduced volume of thalamus is associated with many areas of functioning including the language, motor and executive domains, however, they did not find any significant relationship between thalamus volume and symptoms of patients with schizophreni (Coscia et al., 2009). In another study, Qiu et al. (2009), did not find any relationship between thalamus volumes and

performance of patients on spatial working memory and executive tasks. However, they found significant correlations between these tasks and regionally specific thalamic shape compressions. As can be seen, there is no consistent findings between thalamus volume and problems patients experience, even though there is a significant thalamus volume reduction in patients (Adriano et al., 2010). Due to the importance of information processing between prefrontal cortex (PFC) and thalamus on executive functions and deficits of schizophrenia patients in these functions, Giraldo-Chica et al. (2018) hypothesized a disruption in anatomical connectivity between these structures. Their study showed that, there was a significant relationship between reduced PFC-thalamic connectivity and impaired working memory, however, the same relationship was not found with cognitive flexibility and inhibition. On the contrary to PFC-thalamic connectivity, the connectivity between thalamus and somatosensory and cortical cortices found to be increased in patients.

Frontal lobe functioning has an important role in human cognition as it organizes brain functioning, assist goal-directed and self-regulatory behaviors (Romine & Reynolds, 2005). Planning, suppression of unrelated information, cognitive flexibility, perceptual motor speed, attention shifting, information processing speed are some of the many important functions associate with frontal lobe (Ratti et al., 2002). In patients with schizophrenia, a reduced volume has seen in frontal lobes (Mubarik & Tohid, 2016). Many correlations were found between frontal lobe volume reduction and problems schizophrenia patients are experiencing. In one study, people with schizophrenia and bipolar disorders founded to have reduced inferior frontal lobe which had a significant relationship with their performance on a working memory task (Shepherd et al., 2015). Another study showed reduced bilateral frontal lobe volumes were correlated with high apathy in schizophrenia patients, furthermore, performance of high apathy

group on visuomotor sequencing and verbal learning was significantly lower than low apathy group (Roth et al., 2004). Furthermore, reduced cortical thickness has been reported in patients with schizophrenia (Mubarik & Tohid, 2016), however, cortical thinning suggested to be an effect of antipsychotic medications (Lesh et al., 2015). A relationship might also exist between frontal lobe volumes and positive symptoms of schizophrenia patients. One study showed that there was a negative relationship between hallucination symptoms of patients and grey matter volume within the bilateral frontal and left parietal cortices. The same negative relationship was found between delusion symptoms of patients and frontal cortices grey matter volume (Song et al., 2015). Hirao et al. (2008), investigated the correlation of frontal lobe structures and theory of mind in schizophrenia patients due to its importance in social functioning. Theory of mind is the ability to infer mental state of own and other people, which has been shown to be impaired in people with schizophrenia (Brüne, 2005). Their results suggested that prefrontal cortical reductions in patients' brain might be the underlying pathology behind theory of mind impairments (Hirao et al., 2008).

Hippocampus is a region in the brain associated with many important functions including memory, attention, emotion, olfaction, and navigation (Deshmukh & Knierim, 2012). A meta-analysis of 44 studies showed that patients with schizophrenia had significant hippocampus volume reduction and this reduction is seen in both in first-episode patients and chronic patients (Adriano et al., 2012). One study investigated both volume and shape of hippocampus in schizophrenia patients. Participants in the study were drug-naïve schizophrenia patients, thus the effects of antipsychotic drugs in the brain regions were not a confounding factor. The grey matter volume reduction found in anterior subdivision of hippocampus in patients. Moreover, there was a significant relationship between anterior hippocampus deficit and positive symptoms of

patients. their shape analysis also detected an inward deformation of bilateral hippocampal surface (Kalmady et al., 2017). Another research was done to understand the relationship between hippocampus volume and memory in schizophrenia. They found there was a positive correlation between hippocampus volumes and logical memory in healthy controls, whereas in patients with schizophrenia, there was a positive correlation between posterior hippocampus volume and performance on visual reproduction task. Furthermore, a negative relationship is detected between right anterior hippocampus volume and memory performance (Thoma et al., 2009). Since hippocampus volume reduction is seen both in first-episode and chronic patients, Adriano and colleagues (2012), hypothesized this phenomena may have a neurodevelopmental origin.

How would it be if we did not experience any fear? Although it may sound tempting for a second, fear is a necessary emotion we need to experience to avoid danger. A woman known as S.M., cannot perceive or experience fear. She had a rare disease called Urbach-Wiethe disease which cause calcification of neurons in amygdala (Barrett et al., 2018). Amygdala is the region that is most associated with fear, however, its functions are not limited to that, amygdala also have roles in sexual behavior, emotional processing, reward learning, and motivation (LeDoux, 2007). Niu et al. (2004), found that reduction in amygdala volume in patients with schizophrenia. Furthermore, they found a gender difference that male patients showed reduced bilateral amygdala whereas right amygdala reduction is seen with female patients. For male patients, an asymmetry of left and right volume is also detected. However, there is mixed evidence since not every study replicated this finding. In one study, although some significant structural differences are detected in amygdala of patients, raw amygdala volume did not differ between patients and healthy controls (Rich et al., 2016). Furthermore, a post-mortem study did not detect any

significant difference in amygdala volume in patients (Chance, 2002). However, interesting relationships were found between amygdala volume and problems experienced by schizophrenia patients. Left amygdala volume was found to be predictive of verbal memory performance in schizophrenia and bipolar disorder patient groups, even though such relationship did not exist in healthy controls (Killgore et al., 2009). Right amygdala volume, on the other hand, found to be correlated with suicidal behaviors in patients with schizophrenia (Spoletini et al., 2011). More studies should be done to understand exact structural abnormalities in patients. Mentioned two studies which did not detect any difference in amygdala volume, and this might be related to small sample sizes (Chance et al., 2002; Rich et al., 2016). Wright et al., (2000), included 58 studies in their meta-analysis that collectively had 1588 patients with schizophrenia and detected a reduction in amygdala in patients.

As stated, meta-analysis results showed 1% reduction in white matter in patients (Wright et al., 2000). In their study, White et al. (2011), examined white matter differences between patients and healthy controls by their fractional anisotropy values, which is a common measure of fiber integrity (Leow et al., 2009). The results of the experiment showed lower fractional anisotropy (FA) in patients, especially in frontal, parietal, occipital, and temporal lobes. However, this difference was not found in first-episode patients. Thus, they suggested white matter microstructures are susceptible to progressive alterations (White et al., 2011). Another study found that the existing white matter pathology is associated to neuroinflammation in patients (Najjar & Pearlman, 2015).

Additional to the stated differences, cortical morphology of people with schizophrenia was stated to differ. A complex morphological structure exists for human brain which consists cortical surfaces that are folded or smooth (Matsuda & Ohi, 2018). Surface based morphometry

studies showed that there are some differences between people with schizophrenia and healthy controls. For example, gyrification index (GI), ratio of inner to outer cortical contours, was found to be reduced in patients with schizophrenia (Kulynych et al., 1997). Furthermore, people with schizophrenia showed reduced cortical thickness and increased sulcal width (Janssen et al., 2014). These differences were also correlated with some behavioral measures. For example, numerous correlations between gyrification index and symptoms, and cognitive functions of people with schizophrenia (for a review see Matsuda & Ohi, 2018).

1.3.2 Functional Brain Abnormalities in Schizophrenia

Many studies showed there are some functional abnormalities in patients with schizophrenia. In one study, Venkataraman et al. (2012), investigated resting state functional connectivity of patients with schizophrenia. In comparison with healthy control group, schizophrenia patients found to demonstrate an increased connectivity between parietal and frontal regions and a decreased connectivity between parietal and temporal regions, and between the temporal cortices bilaterally. They also stated significant relationships between positive symptoms of patients and decreased parieto-temporal connectivity. Lastly, they stated a significant relationship between general and negative symptoms and increased fronto-parietal connectivity. Another study found altered amplitude of low frequency fluctuations in patients in default mode network and frontoparietal network (Ren et al., 2013). Additional research was done to find associations between functional abnormalities, and cognitive and behavioral problems of patients. As cognitive control problems of patients found to be related to prefrontal cortex dysfunction, a group of researchers designed an experiment in which participants had to follow the given rules to guide stimulus response mappings while EEG data was acquired. For the task, participant had to press implied buttons in the congruent or incongruent direction of the

stimulus presented. Gamma power differences detected at patient group during delay period of the tasks, gamma cortical oscillatory activity that is associated with problematic cognitive control, found to be present in first episode patients (Minzenberg et al., 2010). Since suicide is highly prevalent among schizophrenia patients, one experiment was done with patients to find the functional correlates of suicidal ideation and suicidal behavior. Results showed that past suicidal ideation was correlated with lower activation in prefrontal cortex during goal representation. Suicidal behavior, on the other hand, was associated with lower activation in premotor cortex (Minzenberg et al., 2014).

1.3.3 Neurochemistry in Schizophrenia

Literature shows the importance of chemical imbalances in schizophrenia. Mainly dopamine, glutamate, GABA, serotonin and oxytocin found to be responsible for the disorder (Bansal, & Chartterjee, 2021). Glutamate and dopamine hypothesis are one of the most longstanding hypotheses of schizophrenia (Howes et al., 2015). Dopaminergic system has critical roles in motor control, motivation, and cognitive function (Klein et al., 2019). The predominant hypothesis of dopamine alterations in schizophrenia is that there is a hyperactivity at dopamine transmission at D2 receptors in limbic striatum (Kegeles et al., 2010) and there is a hypoactivity of dopamine in dorsolateral prefrontal cortex (Toda et al., 2007). The hyperactivity is associated with positive symptoms in schizophrenia (Toda et al., 2007) whereas the hypoactivity of dopamine is associated with negative and cognitive symptoms of schizophrenia (Howes et al., 2017; Toda et al., 2007).

Glutamate, the primary excitatory neurotransmitter, has been shown to have role in schizophrenia (Moghaddam & Javitt, 2012). Glutamate is vital considering it has roles in cognition, memory, movement, behavior, and sensation, and also formation of neural networks

(Sundaram et al., 2012). NMDA receptors are ion channels that are gated simultaneously by voltage, glutamate, and glycine (Moghaddam, 2003). Hypoactivity of glutamate receptors was suggested to have a role in schizophrenia since phencyclidine, that blocks NMDA receptors, is a psychotogen (Seeman, 2009). Even a single dose of NDMA antagonists were shown to be enough to create schizophrenia-like symptoms (Moghaddam, 2003). Glutamate has been associated with some cognitive problems such as working memory in patients. Alterations in glutamate neurotransmission in dorsolateral prefrontal cortex (DLPFC), which mediates working memory, was seen in patients with schizophrenia (Lewis & Moghaddam, 2006). However, there is mixed evidence regarding the relationship between cognitive symptoms and schizophrenia. In their meta-analysis, Iwata et al. (2015), investigated if glutamate positive modulators are effective for cognitive symptoms patients experience, the results showed that glutamate positive modulators were not significantly better than placebo. These mixed results suggest that we need more research to reveal the real relationship between glutamate and schizophrenia.

1.3 Diagnosis of Schizophrenia

Although, treatment is partially effective, the early detection of the disease and early intervention is very important to prevent the worst outcomes in schizophrenia (Razzouk et al., 2006). Most used classification systems for diagnosis are the Diagnostic and Statistical Manual of Mental Disorders (DSM) and the International Classification of Diseases (ICD) (Sampogna et al., 2020).

As stated earlier, in the latest edition of DSM, individual should be diagnosed with schizophrenia if they show at least two of the following: delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior and negative symptoms, for a significant portion of time during 1-month period. At least one of the two of them should be delusions,

hallucinations, or disorganized speech. DSM states level of functioning should be disturbed by the symptoms for at least 6 months (APA,2013). In ICD 11, symptoms of the disorder stated to be delusions, hallucinations, disorganized thinking (formal thought disorder), experiences of influence, passivity or control, negative symptoms, grossly disorganized behavior and psychomotor disturbances. At least two symptoms should be present, and one of which should be delusions, hallucinations, disorganized thinking (formal thought disorder), and experiences of influence, passivity, or control for a diagnosis with schizophrenia. Unlike DSM 5, showing disturbances for 1 month is enough to be diagnosed for ICD 11 (World Health Organization, 2019). Patients, generally get diagnosis through a psychiatric evaluation, which relies on symptoms, medical history, interview, and observation, and this create problems due to heterogeneity of clinical manifestations, which results in less agreement between expert observers (Sorkin et al., 2006). An example can be the research made by Bell et al. (1998), they had a big sample 479 patients who had their first episodes. Their diagnosis was made using 11 different diagnostic systems and the agreement between those systems were 1.7% for the presence and 4.6% for the absence of schizophrenia. Absence of an objective tool for diagnosis is problematic since, as stated, early intervention is very important to prevent the worst outcomes in schizophrenia (Razzouk et al., 2006). One of the main aims of this research is to create an objective tool for the diagnosis of schizophrenia.

1.4 Treatment and Recovery

Some pharmacological and psychosocial interventions is offered for people with schizophrenia (Hasan et al., 2015). First wave of pharmacological treatment was first-generation antipsychotics and clozapine, they are dopamine D2 receptor antagonists, and they are associated with some side effects such as parkinsonism, motor restlessness, and tardive dyskinesia (Goff,

2021). Clozapine is one of the first effective antipsychotics which did not have extrapyramidal adverse effects (Baldessarini, & Frankenburg, 1991). It was seen as a major breakthrough until the discovery of a major side effect, agranulocytosis (Idanpaan-Heikkila et al., 1977; Siskind et al., 2016), which is a life-threatening condition that is characterised by reduction in neutrophil count and vulnerability to infection (Mijovic & Maccabe, 2020). However, despite the side effects it was reintroduced and in their meta-analysis, Siskind et al. (2016) found clozapine was more effective for treatment-refractory schizophrenia.

Second-generation antipsychotics, which combines serotonin 5HT2A and dopamine D2 blockade followed first-generation antipsychotics and found to be effective (Solmi et al., 2017). Goff (2021) summarized this development with the start of risperidone and olanzapine which is analogue to clozapine. Later, starting with aripiprazole, D2/D3 receptor partial agonist drugs started to be used.

One meta-analytic research including 56 studies showed that the long-term use of the antipsychotics demonstrated intermediate efficacy for relapse prevention. Among them, olanzapine was found to be more effective (Zhao et al., 2016). However, some of the patients who respond the treatment are still not independent in their daily lives mostly because drugs are more effective controlling positive symptoms, whereas they are not that effective with negative or general symptoms of schizophrenia (Goff, 2021). Fervaha et al. (2014), suggested negative symptoms such as amotivation and cognitive impairments play a central role in functional outcomes of patients with schizophrenia. Antipsychotics are most effective when it comes to the positive symptoms of patients, but they show reduced effects for negative symptoms and cognitive impairments of patients (Mueser et al., 2013). John Nash, a schizophrenia patient who holds a Nobel Prize on economics stated, ‘I would not treat myself as recovered if I could not

produce good things in my work', reminding the psychiatry community to treatment for schizophrenia should extent aiming reducing positive symptoms of patients (Uno & Coyle, 2019). In their meta-analysis, Correll et al. (2018), showed that early intervention programs which combine psychosocial treatments with antipsychotics showed a significant superiority in terms of outcomes when compared with usual treatments. Cognitive behavioral therapy, assertive community treatments, family psychoeducation, and social skill training are few examples of psychosocial interventions that can be used for patients (Goff, 2021). Cognitive behavioral therapy and social skill training found to be most effective for negative symptoms and the effectivity of CBT was shown to be maintained in the following 6 months after treatment (Elis et al., 2013).

Unfortunately, current success of treating patients with schizophrenia, in general, does not create a very optimistic picture. One big meta-analysis including 50 studies showed that only 13.5% people with schizophrenia and related psychoses met the criteria of recovery. Furthermore, the results shows that there is no proof of better treatment outcomes over the years (Jääskeläinen et al., 2013).

1.5 Heterogeneity of Schizophrenia

Understanding why patients do not respond to treatment is a complicated question to answer due to heterogenous and multidimensional nature of schizophrenia (Kane et al., 2019). Indeed, patients with schizophrenia show clinical and biological heterogeneity, and assuming one average patient group could be the reason of problematic prediction for treatment and outcome (Wolfers et al., 2018). Absence of a valid subtyping scheme can be the reason why traditional efforts could not delineate clinical heterogeneity of schizophrenia (Ahmed et al., 2018).

Kay et al. (1987), created the PANSS scale, a scale for schizophrenia which includes three psychopathological symptom patterns. These factors are positive, negative and general psychopathology. Earlier into this distinction, Crow (1980), labelled positive symptoms, which are presence of abnormal psychological features such as hallucinations, delusions and thought disorder as type 1 syndrome and the diminished or absence of normal functioning such as flattening affect as type 2 syndrome of schizophrenia. In PANSS scale an additional factor called general psychopathology is included in the scale. Kay et al. (1987), included this factor because they consider these symptoms as an important addition to positive/negative subscales. General psychopathology subscale was not added as a completely distinct subscale, but it was included because it can be a parallel measure of severity for schizophrenia. They are considered as non-specific symptoms which have role in the severity of the illness, thus, functioning of the patients. Table 1 list items from PANSS scale, symptom patterns and related items can be seen. PANSS scale has 30 items, and the evaluation takes place as the patients get a score from 1 to 7 on each item, 1 referring to the absence and 7 referring to the extreme level of psychopathology. PANSS scale can be used for computing illness severity, monitoring the progress or to understand from which psychopathological symptom pattern patient suffers the most.

Table 1.

Items and their related psychopathological symptom patterns from PANSS scale (Kay et al., 1987).

Positive subscale	Negative subscale	Generel psychopathology subscale
Delusions	Blunted affect	Somatic concern
Conceptual disorganization	Emotional withdrawal	Anxiety

Hallucinatory behavior	Poor rapport	Guilt feelings
Excitement	Passive-apathetic social	Tension
Grandiosity	withdrawal	Mannerisms & posturing
Suspiciousness	Difficulty in abstract thinking	Depression
Hostility	Lack of spontaneity & flow of	Motor retardation
	conversation	Uncooperativeness
	Stereotyped thinking	Unusual thought content
		Disorientation
		Poor attention
		Lack of judgment & insight
		Disturbance of volition
		Poor impulse control
		Preoccupation
		Active social avoidance

Since one of the most robust biological markers of schizophrenia are the differences in the brain structure, subgrouping patients according to their phenotypic variation will make it possible to create more homogenous groups, which in turn will enable us the detect underlying biological markers without being the victim of heterogeneity (Nenadic et al., 2010). Koutsouleris et al. (2008), investigated the relationship between subtypes of schizophrenia and structural brain abnormalities. They used PANSS scale to measure patients subtype scores, patients were then divided into three groups according to which subscale they had the highest score. For example, if a patient shows more positive symptoms compared to negative and general psychopathology,

they were labelled as patients with predominant positive symptoms. Their results showed a few associations. They found that showing predominantly general psychopathology symptoms was related to bilateral differentiations in temporal, insular and medial prefrontal cortices, predominant positive subscale group showed alterations in left perisylvian regions and reduced thalamic grey matter volume. Lastly, patients with predominantly negative symptoms showed greater alterations compared to other groups including orbitofrontal, medial prefrontal, lateral prefrontal and temporal cortices, and limbic and subcortical structures. However, the search for consistent subgroups for schizophrenia remain inconclusive. Many researchers claimed 3 subscales of PANSS were not adequate to capture the symptomatology, therefore, they offered different factorization solutions. These solutions were inconsistent as researchers offered 4 to 7 factor solutions for PANSS scales (Emsley et al., 2003; Van den Oord et al., 2006; Kim et al., 2012, Levine, & Rabonowitz, 2007; Wallwork et al., 2012; Chen et al., 2020). For example, Chen et al. (2020) stated 4 factors emerged from PANSS scale as positive, negative, affective, and cognitive symptoms. However, sample's neuroimaging data revealed 2 schizophrenia subtypes containing predominantly positive and predominantly negative symptoms.

To summarize, due to heterogenous nature of schizophrenia (1) instead of creating an average profile for patients (Kane et al., 2019)., we need to consider patients individually, and try to find patients' need on the individual level. (2) to better understand the underlying brain structure of schizophrenia, we need to find stable and consistent subtypes, create homogenous groups to find the relationship between brain and subtypes of schizophrenia better (Nenadic et al., 2010). One may wish there was a statistical way to accomplish these aims, luckily there is.

1.6 Machine Learning

One of the most exciting technologies in Artificial Intelligence (AI), machine learning is the field of study that gives computers ability to learn without being explicitly programmed (Das et al., 2015). Machine learning attracts attention due to its ability to accurately predict complex phenomena and producing knowledge about complex relationships in big data (Murdoch et al., 2019). Machine learning is a promising tool to account for the heterogeneity in schizophrenia which results in a more precise knowledge on the etiology of the disease and making predictions for patients individually (Schnack, 2017). The main aim of current study is to create objective tools to make various predictions for schizophrenia patients using their structural brain imaging variables and machine learning methods. This is not the first study which used neuroimaging and machine learning for such purpose. A recent review including 18 studies showed that, machine learning methods can accurately predict if an individual has schizophrenia or not from their brain images. The accuracy levels differed between 63.2% to 93.3% (Winterburn et al., 2019). As for diagnostic predictions, machine learning can be used for prognostic predictions as well (Schnack, 2017). Gong et al. (2020), using a machine learning method and multi-parametric MRI, were able to accurately predict if individual patients will respond to electroconvulsive therapy (ECT) or not. Cui et al. (2019) were able to predict if patients will respond to treatments or not with around 70% accuracy. They used patients' baseline regional activity to make these predictions. Aim of these research efforts is clinical applications, however, it was not possible so far due to some methodological issues. Several problems stated by First et al. (2012) are some diagnostic classification studies has a good overall accuracy but there is a difference of accuracy levels between two groups. Some models show good performance on detecting healthy controls but not patients with schizophrenia. Methodology can be biased in many ways due to pre-processing stages. Furthermore, many studies have small sample sizes. This phenomenon also can be seen

from the stated meta-analysis by Winterburn et al. (2019) as many of the studies discussed there had small sample sizes. Machine learning analysis that is done with small sample, although they can show higher accuracy, their generalizability tends to be low, whereas analysis that is done with big samples tend to show a lower accuracy but a higher generalizability (Schnack & Kahn, 2016). In order to build accurate and reliable models, large sample sizes are needed (Schnack, 2017). Current study, by using a large sample and longitudinal data, aims to create objective tools that are generalizable. A question here waits to be answered. How can machine learning predict if a person has schizophrenia or if a patient will respond to treatment from their brain images?

1.6.1 Classification with Machine Learning

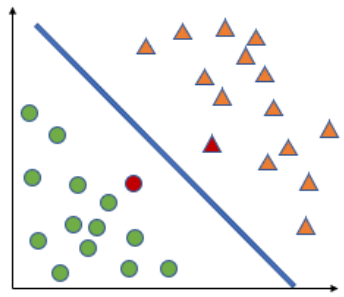
In this section two types of machine learning that are relevant for current study will be discussed, these are supervised and unsupervised learning. Singh and colleagues (2016), in their article, explained how supervised learning works. As implied, supervised learning is learning from experience. In supervised learning, machine is provided with 2 subsets of data, one of them is called train set and the other one is called test set. During training, the learner (machine) finds patterns common to each class from the train set. Then the created model is tested with unseen data, and accordingly to the patterns and features found for each class, model decides which class each test data belongs. The performance and the accuracy of the model is measured according to its classification power. On the other hand, unsupervised machine learning techniques are used when the data is unlabelled. The aim of these techniques is to discover hidden patterns in the data (Saravanan & Sujatha, 2018). In the next section some common supervised and unsupervised learning methods will be discussed.

1.6.2 Support Vector Machines

Support vector machines (SVM) is a machine learning method originally developed by Vapnik and Chervonenkis in 1964. It is a very simple and effective algorithm commonly used for classification and pattern recognition (Jain et al., 2020) Even though, support vector machines can be used for complex classification problems, to illustrate how the linear support vector machine models work, a 2-dimensional 2 class example will be used. In figure 1, there is a two-class problem where we need to separate them according to their shape. The aim of SVM is to obtain a function that creates a hyperplane (blue line) (Jain et al., 2020), in a way that this hyperplane will maximize the space, i.e. margin, in between classes (Nieuwenhuis et al., 2012). Support vector machines got its name from the support vectors (red circle and triangle) which are the nearest data points of each class to the optimal hyperplane (Souli & Lachiri, 2018).

Figure 1

Support vectors and hyperplanes



Note. A separation problem and representation of how SVM works in a 2-dimensional space. Red circle and triangles are support vectors and the line between them is the optimal hyperplane.

Nieuwenhuis et al. (2012) explained how SVM can be used as a diagnostic tool using neuroimaging. Neuroimaging data of people is represented by features congregated into a vector $\mathbf{x}_i$ per participant. A high dimensional feature space contains these vectors, where a decision surface is constructed to separate patients from healthy controls. A decision function is used to create this:

$$y(x_i) = w^T \cdot x_i - b,$$

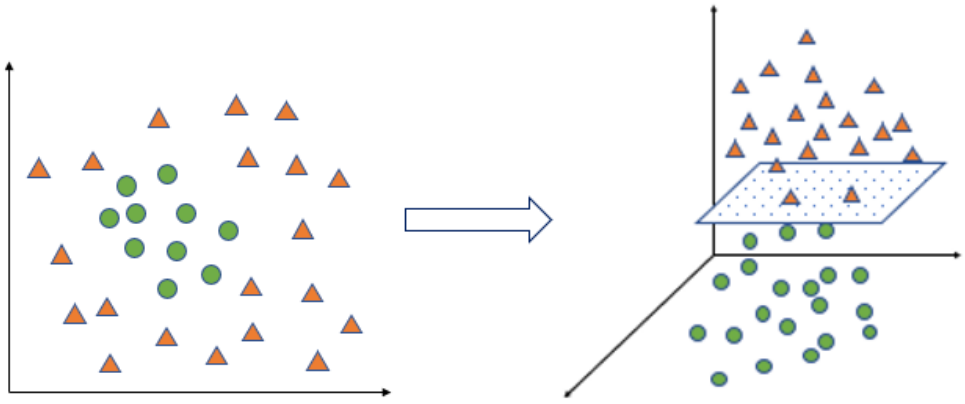
Here, w is a normal vector to this surface, b is an offset. Every participant has a label $\mathbf{t_i}$ indicating if they have schizophrenia or not (patients 1; controls -1), the function is optimized by requiring y(xi)<0 if the label ($\mathbf{t_i}$) is -1 and y(xi)>0 if the label (ti) is +1. In the testing step, participants are classified according to this decision function. Error of the model is calculated through the distance of the subjects to the optimal hyperplane. If the participant classified correctly by the algorithm, error will be 0.

In linear SVM, narrowness of the margin is controlled by the parameter C. penalty C is multiplied by the error per participant, if C is larger, the penalty for wrongly classified participants will be higher and margin will be smaller (Nieuwenhuis et al., 2012).

Dividing groups into two with a straight-line seems easy, but what happens if the data is not separable linearly? In these cases, kernel functions can be used to map the data into a different higher dimensional space where a separation with a hyperplane is possible (Bhavsar & Panchal, 2012). Figure 2 can be illustrative on how this phenomenon works.

Figure 2

Kernel trick



Note. Transformation of linearly inseparable data into a higher dimensional space.

The most common types of kernels are radial, polynomial, and sigmoid kernel functions. Radial kernel (RBF) is the most used function for SVM classification, it is a similarity function that involves Euclidean distance calculation. The general formula is the following formula (Ghosh et al., 2019):

$$K(x, x') = exp\,(-\gamma\|x - x'\|^2)$$

Furthermore, the formula for γ is:

$$\gamma = \frac{1}{2\sigma^2}$$

RBF-SVM has two parameters that needs to be set before analysis. These are the C, cost, also needs to be set for linear SVM, and σ (sigma) (Yang et al., 2010). Sigma specifies the width of kernel function, if it is high, the region of decision becomes broad and it leads to a smoother decision boundary, if the opposite is the case, decision boundary becomes high and around the data points, islands of decision boundaries can be seen (Al-Mejibli et al., 2020). In short, sigma decides how flexible the classifier will be. (Ben-Hur et al., 2008):

Support vector machine classification with polynomial kernel, functions as the similarity of the samples is represented by polynomials of original variables in a feature space (Ghosh et al., 2019). Equation of the function is:

$$K(\mathbf{x}_i, \mathbf{x}_j) = (\gamma \mathbf{x}_i^T \mathbf{x}_j + r)^d, \gamma > 0$$

Width of the kernel function is γ and *d* is used for polynomial degree term (Srivastava et al., 2012). Degree is related to flexibility of the classifier. Lowest degree polynomial, it is when the degree is 1, is a linear kernel, as degree gets higher values, decision boundaries become more flexible (Ben-Hur et al., 2008).

Lastly, sigmoid kernel, is a function inspired by neural networks. It is equivalent to a two-layer neural network (Fadel et al., 2016) which have three hyperparameters c (cost), ϒ (gamma) and r, which controls the threshold for mapping (Ab Kadera et al., 2019). Equation is given below(Grama et al., 2017):

$$K(\mathbf{x}_i, \mathbf{x}_j) = \tanh(\gamma \mathbf{x}_i^T \mathbf{x}_j + r),\ \gamma > 0$$

Support vector machines were widely used for schizophrenia research. Winterburn and colleagues, in their reviews, used 18 diagnostic classification studies for schizophrenia and 7 of the studies used support vector machines (Borgwardt et al, 2013; Davatzikos et al., 2005; Fan et al., 2007; Nieuwenhuis et al., 2012; Pettersson-Yeo et al., 2013; Yushkevich et al., 2005; Zanetti et al., 2013).

1.6.3 Random Forests

Random forest is, as presented by Breiman (2001), an ensemble of decision tree classifiers that includes multiple randomizations. Random subsets of the training data are used to grow each tree. Each decision node is constructed by random subsets of features (such as white matter volume). In random forest, an importance score called Gini Importance is calculated for each feature showing its discriminative power (Venkataraman et al., 2012). To illustrate, let's imagine we are interested in classifying patients with schizophrenia into two classes according to their illness severity; less and more severe patients, with their structural brain imaging data. If we look at the Gini importance scores of each feature, we can understand which features have more discriminative power when it comes to differentiating less and more severe patients. If grey matter volume has a higher gini importance score than white matter volume, it will mean that grey matter volume is a better variable to check for such classification. This randomization process is not stopped until each leaf of tree defines unique class which would be, according to

above example, severity of the illness. The final decision in random forest is given through a majority voting among decision trees (Venkataraman et al., 2012).

Many researchers used random forest to discover the relationship between schizophrenia and the brain. One example is a study by Greenstein et al. (2012), they used random forest to classify childhood onset schizophrenia patients and healthy controls using their neuroimaging data. They achieved 73.7% accuracy with their analysis. Another study by Talpalaru et al., (2019), used many methods including random forests to classify patients with schizophrenia into 3 groups: patients with low and high symptom burden, and predominantly positive symptom burden. They reached the best accuracy with random forests. This outperforming could be due to randomization and majority voting features random forest has.

1.6.4 Linear Regression

One of the most used and simplest machine learning techniques is linear regression(Maulud & Abdulazeez, 2020). In regression analysis a dependent variable value "y" is found based on independent value(s) x1,x2…xk (Roopa & Asha, 2019). When there is only one independent variable, simple linear regression is used and it can be formulated as "$y = \beta_0 + \beta_1 x + \varepsilon$" where β_0 is the intercept, β_1 is the slope of the regression line and ε is the error (Zou et al., 2003; Maulud & Abdulazeez, 2020). When there are multiple independent variables, it is called multivariate linear regression which can be similarly formulated as $y = \beta_0 + \beta_1 x_1 + \cdots + \beta_m x_m + \varepsilon$.

For two class problems, linear regression would create unsatisfactory results. In these cases, another regression model called logistic regression can be used (LaValley, 2008). The calculation of logistic regression is not that different from linear regression. Assuming the

logistic regression analysis will be used for multiple independent variables, the equation can be stated as:

$$\ln(\hat{Y}/1-\hat{Y}) = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \ldots \beta_i X_i$$

The categorization problem is solved by mathematically transferring linear regression equation to yield natural log of the odds. Here $\hat{Y}$ refers to one class and $1-\hat{Y}$ refers the other (Stoltzfus, 2011).

Regression analyses were used in many schizophrenia studies. One example is the study by Hulshoff Pol et al. (2002), where they used linear regression to discover the relationship between brain volumes of schizophrenia patients and their age. A steeper regression slope between age and grey matter volume of patients compared to healthy controls was stated, which means age, compared to healthy controls, was correlated with greater reduction in grey matter volume of patients.

Another study by Leucht et al. (2007), used logistic regression to investigate the relationship between demographic variables of schizophrenia patients, their symptoms, and their treatment response. Higher age, more symptoms on baseline and longer illness duration was associated with nonresponse to treatment.

1.6.5 XGBoost

XGBoost (extreme gradient boosting)(Sheridan et al., 2016), a scalable tree boosting system, is a commonly used and highly effective machine learning method which achieves state-of-art results (Chen & Guestrin, 2016). Boosting is the method that refers to running a learning algorithm many times in different distributions using the training data and then merging the classifiers with the aim of a better performance(Freund & Schapire, 1996). Chen & Guestrin (2016) stated, XGBoost is the common ML method that is used by the winning team in machine

learning competitions. Furthermore, it is 10 times faster due to parallel and distributed computation. The main idea behind the algorithm is that it builds classification trees individually, so each new tree uses the residuals of the previous one. So, an outcome is predicted as the new model corrects the previous errors (Pesantez-Narvaez et al., 2019). Nielsen (2016) explained the reasons why XGBoost is very effective compared to other machine learning methods. XGBoost, due to being a tree boosting method, is very effective because of its rich representational ability. It uses an interesting penalization system which gives the opportunity to model to have many numbers of terminal nodes. Lastly, extra randomization parameter is used in XGBoost.

1.6.6 K-means Clustering and Factor Analysis

Until now, relevant supervised machine learning methods are discussed. Unsupervised machine learning techniques are used when the data is unlabelled. The aim of these techniques is to discover hidden patterns in the data (Saravanan & Sujatha, 2018). Two common means of unsupervised learning are clustering and dimensionality reduction (Ghahramani, 2004). Two methods, which are relevant to this paper will be discussed in this section. First of them is a clustering method called K-means clustering. Clustering is simply finding homogenous groups of data points in a data set. These homogenous groups are called clusters, and, in these clusters, density is locally higher compared to other regions (Likas et al., 2003). In their paper, Kodinariya and Makwana (2013) explained how k-means clustering works. The number of the clusters should be determined before K-means clustering. This algorithm defines k centroids, one for each cluster. One of the aims of the algorithm to define these centroids as far away as possible from each other. Later each data point is assigned to the nearest centroid, so these data points will be under that cluster. When every point is assigned, re-calculation of centroids will take place and algorithm will make corrections on point assignments to the nearest centroids.

This process will end when there is no need for reassignment and every data point is under a cluster. One example of usage of K-means clustering in schizophrenia research is Bell and colleagues (2010) where they investigated if three memory profiles of patients, nearly normal, subcortical impairment, and cortical impairment, could be confirmed by K-means clustering and the results showed that, indeed, K-means supported these three memory profiles for patients.

Real world data generally has a high dimensionality, to adequate handling of such data may require a meaningful reduction of this dimensionality (van der Maaten et al., 2009). An example of why this could be necessary was stated in the paper earlier. Due to heterogenous nature of schizophrenia, creating homogenous patient subgroups may help us to understand symptomatology and their underlying brain abnormalities better (Nenadic et al., 2010).

Factor analysis is a dimensionality reduction method (Khosla, 2004). The initial information is grouped in factor analysis, according to their correlation between variables. The main aim is to create groups that are strongly correlated between each other and less correlated with other groups, each groups represent a factor (Tăutan et al., 2021).

1.7 Related Work

In this section a brief summary of important points, research questions and hypothesis will be given.

Schizophrenia is a disabling mental health issue (Mueser & Jeste, 2008) that effects around 1% of the population in the world (Jones et al., 2005). It is associated with 10-25 years decrease in life expectancy (Laursen et al., 2012) and treatment does not seem to be really effective as only %13.5 of people with schizophrenia meet the recovery criteria (Jääskeläinen et al., 2013). Thus, there is a significant need for schizophrenia research to increase the life quality of patients.

Early detection and early intervention of the disorder is important to prevent worst outcomes (Razzouk et al., 2006). Early detection is problematic due to absence of an objective tool for diagnosis (Bell et al., 1998). Furthermore, treatment failure could be due to the assumption of one average patient group, although, patients with schizophrenia show clinical and biological heterogeneity (Wolfers et al., 2018). Machine learning is a promising tool to make predictions about patients individually (Schnack, 2017). Thus, first aim of this study is to create an objective diagnostic tool for schizophrenia which is accurate and generalizable. This tool was used to make predictions about diagnosis according to the structural brain imaging data of people, since structural brain abnormalities are one of the most robust biological markers of schizophrenia (Nenadic et al., 2010). Here is the first research question of the study:

Research question 1: Is it possible to accurately predict if people have schizophrenia using their brain images and supervised machine learning methods?

Hypothesis 1: Supervised machine learning methods can accurately predict if people have schizophrenia using their structural brain imaging variables.

Due to the importance of finding subgroups among patients that are homogenous (Nenadic et al., 2010), another aim of this study is to see if it is possible to classify patients into subgroups using their symptom pattern data and their brain images. An illustration can be given for why this is an important aim. Let us think about two schizophrenia patients, patient A and patient B. According to DSM-5, to get a diagnosis, a person should show at least two of the following: delusions, hallucinations, disorganized speech, grossly disorganized or catatonic behavior and negative symptoms (APA, 2013). Imagine patient A shows symptoms of visual hallucinations, delusions, and suffer from depression and anxiety. Imagine patient B shows different symptoms than patient A, and shows disorganization during speech, has a rigid posture,

and has blunted affect. Both patients would get a diagnosis of schizophrenia according to DSM-5 if they show functional disturbances over 6 months (APA, 2013). Searching for biomarkers is an important task, however, assumption of patient A and patient B will show exactly same abnormalities in their brain could be problematic since what they suffer from is different. Furthermore, throughout this paper, heterogeneity mentioned many times as a problem. When it is possible to objectively classify patients to subgroups, it might be easier to detect what kind of treatments work better for which group and it could be possible to raise the success rate for treatments. Measurement of symptoms was done through PANSS scale in this study, and as already mentioned, PANSS scale has three original factors (symptom patterns), positive, negative, and general psychopathology (Kay et al., 1987). In this study, each patient was assigned to one of these three groups accordingly to which symptom pattern they experience the most. Then, machine learning methods were trained using their class information (which symptom pattern they suffer from most) and their structural brain images. Later, these algorithms were tested using data of a group of patients which were not in the trainset. It will be seen if the algorithm can accurately predict which symptom patterns these patients suffer the most only by using their structural brain imaging data. Three patient groups are:

1) Patients with predominant general psychopathology symptoms
2) Patients with predominant negative symptoms
3) Patients with predominant positive symptoms

Related research question and hypothesis are:

Research question 2: Can supervised machine learning methods accurately classify patients into three groups according to their symptom patterns using their structural brain imaging data?

Hypothesis 2: Supervised machine learning methods can accurately classify patients into three groups according to their symptom patterns using their structural brain imaging data.

Although PANSS scale has three original symptom patterns, some researchers who used factor analysis for PANSS scales found different number of factors (Emsley et al., 2003; Van den Oord et al., 2006; Kim et al., 2012, Levine, & Rabonowitz, 2007; Wallwork et al., 2012; Chen et al., 2020), meaning there could be different number of symptom patterns than three. To investigate this issue further, unsupervised machine learning methods were used to first detect different symptom patterns. Then patients were divided into groups according to these new symptom patterns and from which they suffer the most. Later, supervised machine learning methods were used to see if the algorithms can accurately classify patients according to these subgroups using their structural brain imaging data. There is an expectation that this classification outperforms the previously mentioned classification since some researchers found different symptom patterns than three originally proposed symptom patterns for PANSS. More accurate classification through symptom patterns is expected to result in more accurate MRI-based classification. This brings us to the new research question and hypothesis:

Research question 3) Can an unsupervised machine learning method to classify patients increase the accuracy of MRI-based subtype classification?

Hypothesis 3) Classification of patients according to their symptom patterns through an unsupervised machine learning method can increase the accuracy of MRI-based subtype classification.

So far, main aims of this study are stated. However, the dataset used in this study has various information about patients and this makes additional classification possible. Many studies found relationship between illness duration and structural brain abnormalities in patients

(e.g Haijma et al., 2013; Tanskanen et al., 2010). Tanskanen et al. (2010) found that grey and white matter deficits were associated with illness duration. Haijma et al. (2013) stated larger volume reductions were associated with advanced duration of illness. These were specifically shown in prefrontal regions, STG, Heschl's gyrus, and parietal lobe. Thus, it is hypothesized that a classification for illness duration could be possible using brain imaging variables. These studies generally used regression and correlation to investigate the relationship. Up to author's knowledge there is no classification study made for illness duration in schizophrenia.

Illness severity is also founded to be correlated with many brain imaging variables. For example, in their study, Yoshihara et al. (2008), found that more severe positive symptoms were correlated with reduced grey matter volume in posterior cingulate gyrus, and more severe negative symptoms were correlated with grey matter volume in thalamus. Ho (2003) suggested greater negative symptom severity was related to decreased white matter volume in the frontal lobe. Thus, a classification for illness severity was thought to be possible using structural brain imaging variables. The related hypotheses and research questions are:

Research question 4: Can supervised machine learning methods accurately classify patients according to their illness severity using their brain imaging variables?

Hypothesis 4: Supervised machine learning methods can accurately classify patients according to their illness severity using their brain imaging variables

Research question 5: Can supervised machine learning methods accurately classify patients according to their illness duration using their brain imaging variables?

Hypothesis 5: Supervised machine learning methods can accurately classify patients according to their illness duration using their brain imaging variables.

Next section will discuss the methodology of this study.

CHAPTER 2

METHODS

2.1) Participants

Data of 467 participants included in the dataset. 169 of them were schizophrenia patients (SZ group). Among patients, 131 of them were male and 38 of them were female. Age mean of schizophrenia patients at baseline was 29.85 years (sd= 9.26), patient group's age ranged from 16 to 56 years. 298 participants were healthy controls (HC group). Among them 167 of the participants were male and 131 of the participants were female. Age mean of HC group at baseline was 30.42 years (sd=10.96), ranging from 16 to 64 years.

Although 467 participants included in the study, data included 1087 brain images since image acquisition took place more than one time for some participants. Three timepoints exist in the dataset. Timepoint 1 includes brain images of 467 participants, timepoint 2 includes brain images of 461 participants and the timepoint 3 includes brain images of 159 participants. Out of 1087 brain images, 700 of them belonged to HC group and 387 of them belong to the patient group. In timepoint 2, out of 461 brain images, 168 belonged to people with schizophrenia. 130 of them belonged to male patients and 38 of them belonged to female patients. Age mean of patients was 34 years (sd=9.56). 293 brain images at time point 2 belonged to HC group. 163 of them were males and 130 of them were females, age mean of HC group at time point 2 was 34.35 years (sd=11.29). Out of 159 brain images in timepoint 3, 50 of them belonged to SC group. Among them only 7 of them were females. Mean age for patients in time point 3 in years was 34.34 (sd=5.72). 109 of the images belonged to healthy controls and 59 of them were males and 50 of them were females. Age mean of HC group at time point 3 was 34.1 years (sd=8.08).

2.2 Data acquisition

Participants were selected from a large longitudinal sample of two cohorts, Utrecht Schizophrenia project and the Genetic Risk and Outcome of Psychosis (GROUP) consortium (Korver et al., 2012), in Utrecht, the Netherlands. Patient group were recruited in various inpatient and outpatient facilities. Ethical approval was provided by University Medical Center Utrecht. All participants provided written informed consent. The dataset used in this study was owned by Department of Psychiatry, Brain Division, UMC Utrecht.

In the Utrecht Schizophrenia Project, patients had to fulfil diagnostic criteria for schizophrenia according to DSM-IV, and their diagnosis was confirmed 1 year later. The criteria of being healthy control was to never have a diagnosis of a mental disorder.

In the GROUP consortium (Korver et al., 2012), patients again had to fulfil diagnostic criteria for schizophrenia according to DSM-IV. Healthy controls, on the other hand, were selected from people who did not show any psychotic problems in their life and their first- or second-degree relatives did not have any psychotic disorder.

Dataset included information of age at scan in years, sex and IQ values estimated with Wechsler Adult Intelligence Scale (WAIS III) (Wechsler, & Psychological Corporation, 1997). PANSS scale was used to assess clinical severity. α coefficients of the subscales were .73, .83 and .79 for positive, negative, and general psychopathology subscale respectively (Kay et al., 1987). Illness duration was the subtraction of age of illness onset from age at the time of scan. Antipsychotic daily dose was converted into chlorpromazine milligram equivalents per patient.

2.3 Image Acquisition

Two scanners were used, and they had the same field strength, acquisition protocol and vendor. Scanning took place twice on either a Philips Intera or Achieva 1.5 T and a T1-weighted, 3-dimensional, fast-field echo scan with 160-180 1.2 mm contiguous coronal slices (echo time

[TE], 4.6 ms; repetition time [TR], 30 ms; flip angle, 30°; field of view [FOV], 256 mm; in-plane voxel size, 1x1 mm2) was acquired.

2.4 Image Processing

FreeSurfer analysis suite (v5.1) was used to acquire detailed anatomical information for each participant (Fischl, 2012). Total brain volume was calculated as sum of total white and grey matter volumes. Then, extraction from FreeSurfer output took place for cortical volume, and mean global cortical thickness. Morphologic variables were extracted with BrainVISA software (v4.5) using the Morphologist Toolbox and Mindboggle software using default settings (Mangin et al. 2004; Klein et al. 2017). For each fold, sulcal area was the total surface area of the medial sulcal surface, sulcal length was defined as the distance between the median sulcal surface and the hull. Furthermore, sulcal width was measured as the distance between gyral banks, averaged over all points along the median sulcal surface. Kochunov et al. (2009) defined gyral span as the distance 2 points on gyral white matter mesh on either side of the sulcal surface. Thus, gyral span variables were calculated as the extending sulcan span tracings until they intersected with gyral grey-white matter interface. For an illustration of measurement of these images, it is possible to see figure 1B in Diaz-Caneja et al. (2021). Finally, the extracted brain imaging variables were: Total brain volume (cm^3), cortex volume (cm^3), cerebral white matter volume (cm^3), subcortical grey matter volume (cm^3), total grey matter volume (cm^3), intracranial volume (cm^3), cortical thickness (mm), pial surface area (cm^2), sulcal area (cm^2), gyral span (cm), sulcal width, depth, and length (cm). Furthermore, dataset included information of sulcation index and local gyrification index of participants. The sulcation index (SI) is the ratio between total sulcal surface area and brain hull area, on the other hand, gyrification index refer to the measure of cortical folding which is the ratio between the length of outer folded surface and length of outer

surface excluding sulci (Zilles et al., 1988). Higher ratio of both sulcation index and gyrification index signifies increased sulcation.

More information about image acquisition, preprocessing, and participant collection can be found in (Hulshoff Pol et al. 2001; Korver et al. 2012; Kubota et al. 2015; Diaz-Caneja et al., 2021; Janssen et al., 2021).

2.5 Statistical Analysis and Machine Learning

Data management and all statistical analyses were performed using RStudio version 4.1.1 (R Core Team, 2013). Various statistical analysis and machine learning methods were used for this study. In the following section methods used and which libraries used to conduct the analysis will be explained.

2.5.1. Controlling Variables

Significant changes in global and regional brain volumes can be seen with healthy aging (Scahill et al., 2003). Additionally, there is a significant brain volume difference between males and females (Giedd et al., 1996). The real aim of this study to find the structural brain patterns related to schizophrenia, and due to this, the effect of the age and sex needed to be eliminated. Without controlling these variables, we cannot be sure that if the differences come from schizophrenia or there is an effect of age and sex on the results. Thus, age and sex were controlled for brain imaging variables using linear regression. This is a common method used by other studies (e.g., Bansal et al., 2012; Nieuwenhuis et al., 2012),

Residuals are the difference between observed values (the value of the data) and predicted values ($e = y - \hat{y}$) (Gloutney & Clark, 1991). Linear regression analyses were performed using age and sex as independent variables and each brain imaging variables (separately) as the dependent variable. This analysis result with predicted brain imaging values

according to age and sex for each participant. Then, observed brain imaging values were subtracted from predicted ones to eliminate the effects of age and sex for each participant. More details about this technique can be found in Pfefferbaum et al.(1992).

2.5.2 Sociodemographic and Clinical Data

Various statistical analyses were done to explore sociodemographic and clinical data of participants. Independent samples t-tests can be used to investigate the difference of the means of two different groups on a dependent variable (Nasr & Kuasar, 2009). In this study, t-tests were used to explore the differences of means between groups. Specifically, t-tests were used to investigate difference of IQ levels between HC and SZ group. Furthermore, although it was not planned, due to the importance of sulcal width, sulcal width values of HC and SZ group were compared using t-test. For this analysis, 'stats' package in RStudio was used (R Core Team, 2013). Linear regression was used to see the relationship between sex, age and total brain volume, again 'stats' package was used for this analysis.

2.5.3 Classification and Supervised Learning Models

Various supervised machine learning models were used throughout the study. For supervised machine learning classifications, brain imaging data of participants were divided into two for a training and testing set. Later, these methods were trained using the training set and learner, detects the common patterns in brain imaging data for each group. Later the model is tested using test set, a part of dataset that model is not familiar with. Finally, a classification made by the model according to found patterns on training process (Singh et al., 2016).

For diagnostic classification, four different methods were used. These methods were support vector machines with radial kernel, random forests, logistic regression, and XGBoost. For supervised classifications according to clinical pattern of patients, support vector machines

and random forests were used. For classification according to illness duration and severity, support vector machines were used. The details about how these models work can be found in the machine learning section in the introduction.

Support vector machines is a commonly used machine learning method that was found especially successful for datasets with small sample and high dimensionality (Yan et al., 2004). A trial-and-error approach was used for the decision of kernel. Radial kernel was the most successful non-linear kernel type, thus used for classification. 'e1071' package was used for support vector machines (Meyer et al., 2021). For each classification, gamma and cost values were selected according to the maximum reached classification accuracy.

Random forests are a commonly used machine learning method, owing its success to randomization and being an ensemble method (Breiman, 2001; Venkataraman et al., 2012). Random forests calculate a Gini importance score for each feature showing how important each variable was for differentiating between classes (Venkataraman et al., 2012). Thus, these makes random forests the perfect candidate especially when the aim is just not classification but also exploration of the features. For this method, 'randomForest' package was used (Liaw, & Wiener, 2002). For random forest analysis using this package, number of trees were selected using a trial-and-error method until the maximum classification accuracy was reached.

Logistic regression was used only for diagnostic classification. 'stats' package was used for logistic regressions (R Core Team, 2013). In logistic regression problem of classification is solved through transforming regression equation to yield natural log of the odds (Stoltzfus, 2011). Logistic regression, similar to random forests, gives information about how important each feature was for the classification. It is also possible to reach p values for each variable using 'stats' package.

XGBoost (extreme gradient boosting) (Sheridan et al., 2016) is a scalable tree boosting system that is commonly used and achieves state-of-art results (Chen & Guestrin, 2016). XGBoost was also only used for diagnostic classification. 'xgboost' package was used for this method (Chen et al., 2021). XGBoost has many parameters one can specify but only two of those parameters were controlled, which were eta and gamma. Eta parameter controls the learning rate, and gamma parameter controls how conservative the algorithm will be (Chen et al., 2019). These parameters were selected according to classification success of the model.

To summarize, support vector machines were used for every classification, random forests were also used for subtype problem and logistic regression and XGBoost were only used for diagnostic classification. The reason behind this is that diagnostic classification wanted to be evaluated in a greater term. Furthermore, support vector machines are commonly used and showed great success. Along with that, computational cost is small, classification takes little time, and it is not complicated. These made SVM a great candidate model for this study, thus used more often. However, there are also some disadvantages of SVM, although SVM is commonly used for 2-class problems, it is generally avoided when the number of classes are more than 2 (Duan & Keerthi, 2005). The only 3 class problem was the classification for symptom patterns and for that, random forest was used along with SVM.

2.5.4 Unsupervised learning models

Unsupervised learning was used only for one classification problem. As stated, different number of factors than original three factors were found for PANSS scale (Emsley et al., 2003; Van den Oord et al., 2006; Kim et al., 2012, Levine, & Rabonowitz, 2007; Wallwork et al., 2012; Chen et al., 2020). Thus, various analyses were done to explore the clusters in the current dataset (Kodinariya & Makvana, 2013).

First of all, k-means clustering was used to detect the number of clusters using an elbow method (Andrew, 2012). It is a visual method to detect which number of clusters is appropriate for the data considering the cost that comes with the number of clusters (Kodinariya & Makvana, 2013). 'ClusterR' package was used for this purpose (Mouselimis, 2021). Factor analysis is used for creating groups that are correlated between each other, but they have less correlation with other groups (Tăutan et al., 2021). Thus, factor analysis was used to detect the factors for PANSS scale and to see which items are correlated with detected factors. Later, the factors were named, and patients were redivided according to from which symptom pattern (factor) they suffer more. Confirmatory factor analysis (CFA) might be conducted to see the fit of the factor models (Sarmento, & Costa, 2019). CFA was used to see the differences between three and two factor models using 'lavaan' library in RStudio (Rosseel, 2012). Classification according to this grouping was done with support vector machines and random forests.

2.5.5 Evaluation metrics for classifications

Many evaluation metrics were used for this study. For every classification, an overall accuracy was found. Overall accuracy is found by the number of correctly classified brain imaging data divided by number of all brain imaging data (Benba et al., 2017). Figure 3 illustrates how confusion matrix works for a binary classification, figure 4 illustrates an actual confusion matrix from the study. Actual category of participants comes from data, and predicted category is the results of the classification process by the algorithms. For diagnostic classification in this study, true positive refers to the number of brain imaging data that was correctly classified as SZ group. True negative refers to the number of brain imaging data that was correctly classified as HC group. False positive and false negative refers to the number of misclassifications for HC and SZ group respectively.

Figure 3

Confusion matrix for 2 group problems.

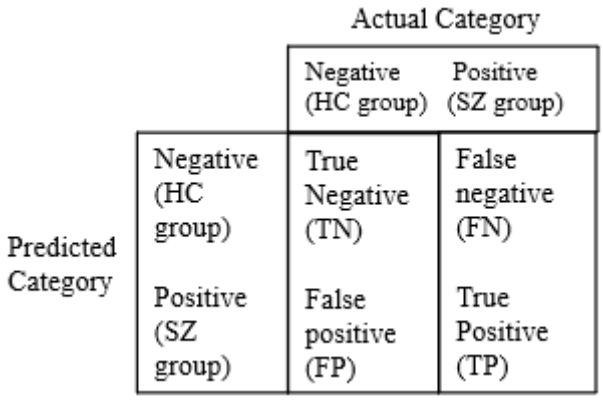
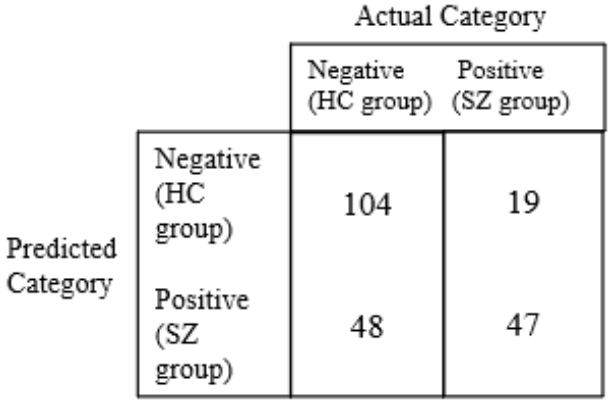


Figure 4

Confusion matrix example from the study. This classification was done for diagnosis using SVM with radial kernel.



The formula for overall accuracy is:

$$\frac{TP+TN}{TP+TN+FP+FN}$$

Simply, accuracy is the amount of data that was correctly classified by the algorithm. Additional to the accuracy, true positive and true negative rates were discovered. True positive rates (TPR), also called sensitivity, is the amount of SZ group data that was correctly classified,

whereas true negative rates (TNR), also called specificity, is the amount of HC group data that was correctly classified (Hong & Oh, 2021). Here is the formula for sensitivity (TPR):

$$\frac{TP}{TP+FN}$$

Here is the formula for specificity (TNR):

$$\frac{TN}{TN+FP}$$

For classifications other than diagnostic classifications, overall accuracy was found, later the amount of correctly classified data was stated for each group similar to diagnostic evaluation metrices.

For each classification, area under curve (AUC) values were also reported. AUC is a useful method of classification (Riquelme et al., 2008), especially when the data is not balanced (Saifudin et al., 2019). As stated, around 65% of the brain imaging data belongs to healthy controls. Thus, AUC is an important evaluation metric for this study. AUC is the area under a ROC curve, which are the curves used to determine the ideal cut-off values for specificity and sensitivity (Fan et al., 2006). Maximum AUC value is 1, which is the perfect classifier, the minimum AUC level is 0.5, which means classifier does not have any discriminative power (Wismüller & Vosoughi, 2021). According to Salgado-Pineda and colleagues (2018), classifier with an AUC level less than 0.6 has no discriminatory power. AUC levels between 0.6 and 0.7 can be considered as poor, AUC levels between 0.7 and 0.8 can be considered fair. If the AUC level is more than 0.8, it can be considered good and if it is more than 0.9, the discriminatory power can be considered excellent. AUC values were found using the ‘pROC’ package of R

(Robin et al., 2011). Graphs were created using the base package of RStudio (R Core Team, 2013) and ggplot2 (Wickham, 2016).

CHAPTER 3

RESULTS

3.1 Sociodemographic and Clinical Findings

As stated in the methods section, dataset includes various information of the participants. In total, 361 of the participants' IQ levels were available in the dataset. Out of these 361, 125 of them were the SZ group and 236 of them were HC group. Average IQ level of participants was 106 and standard deviation(sd) is 17.04. Average IQ level of the SZ group was 97.85 with an sd of 16.47. Mean IQ level of control group was 110.34 and the sd was 15.74. These IQ levels refer to levels in the first scan, baseline, for each participant. T-test comparison showed that IQ level of two groups were significantly different ($t(242.97) = 6.95$, $p < .01$).

Dataset includes PANSS scale scores of patients. PANSS scale has 30 items, on each item, patients get a score from 1 to 7 on each item, 1 referring to the absence and 7 referring to the extreme level of psychopathology (Kay et al. 1987). Mean values of participants on PANSS scale on baseline is presented in table 2.

Table 2

PANSS total and subscale scores of participants on timepoint 1 (baseline)

PANSS scores	n	Mean	Standard deviation
PANSS total score	137	63.34/210	19.26
PANSS positive subscale score	140	15.17/49	5.57
PANSS negative subscale score	140	16.5/49	5.94
PANSS general psychopathology subscale score	137	32.35/112	9.63

Note. n refers to the number of patients whose data were available for corresponding score

116 out of 169 patients were using antipsychotics. Additional to their daily dose of antipsychotics, information on patient's age of onset of the disorder and illness duration on baseline can be found on table 3.

Table 3.

Age of onset, antipsychotics daily dose, and illness duration of patients on timepoint 1 (baseline)

Variables	n	Mean	Standard deviation
Age of onset	110	21.87	5.33
Illness duration	106	6.51	6.62
Antipsychotics daily dose	84	322.08	176.3454

Note. Illness duration and age of onset was stated in years, antipsychotic daily dose was stated in mg in chlorpromazine equivalents.

3.3 Classification for Diagnosis

As age and sex has an effect on the brain, as stated, linear regression was used to eliminate their effects. As expected, sex was a significant predictor of total brain volume (AdjR2=0.278, $F(1,1085)= 420.8$, $p<0.001$). Furthermore, age is also a significant predictor of total brain volume (AdjR2=0.065, $F(1,1085)= 77.44$, $p<0.001$).

First, a classification took place with support vector machines. First step of supervised machine learning is formation of train and test sets (Singh et al., 2016). Caret and data.table packages (Kuhn, 2021; Dowle, & Srinivasan, 2021) were used in RStudio for splitting the data and formation of confusion matrix. As mentioned, the dataset involves data of 1087 brain images, some participants' brain images were acquired more than one time. An important issue here is that for fair comparisons, test set should be independent (Kelly et al., 2019). Due to this

knowledge, data is divided into train and test sets in a way that no participants' brain images will not be included in both tests. Brain imaging data of a participant, even if their brain images were acquired more than one time, ended up in the same subset, either train or test set. First, brain imaging variables of participants that are corrected according to age and sex were shuffled and then they were divided into two in a way that train set would include %80 of the data and test set include %20 of the data. There are no exact guidelines of splitting ratios for train and test sets, however, 80/20 ratio is a common method which was used many times in schizophrenia research (e.g., Chatterjee et al., 2020; Jilka et al., 2022; Tan et al., 2021; Tarchi et al., 2021).

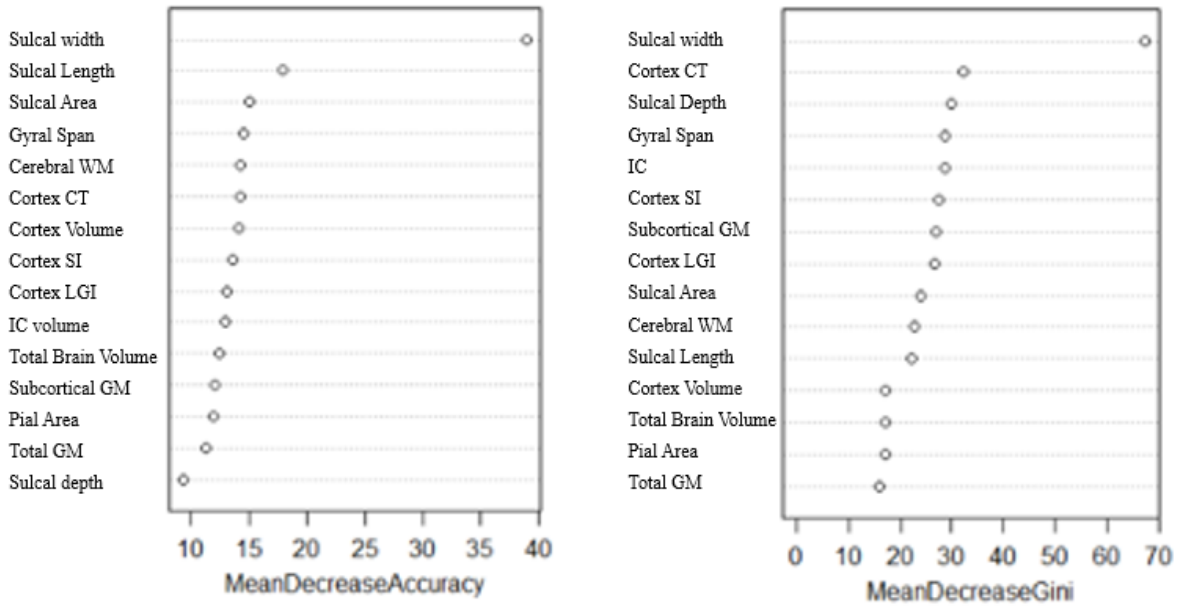
After preparing the test and train set, svm, using the package 'e1071' was conducted. A trial-and-error approach was used for kernel selection and radial kernel, one of the most common types of kernels is selected for the algorithm. For an SVM analysis with radial kernel, sigma and cost parameters needs to be set prior to analysis (Yang et al., 2010). Gamma instead of sigma is required by 'e1071' package, those values are related, and their relation was explained in the introduction section. Value of the parameters selected through trial-and-error. The default values of both cost and gamma in e1071 is 1. Thus, starting with the default values, increasing, and decreasing values of gamma and cost tried until the best accuracy level is reached. The best accuracy level is reached by gamma value of 0.008 and cost value of 0.006. Overall accuracy was 69.2%, 71.2% of the patients were correctly classified, %68.4 of HC group was correctly classified (sensitivity=0.71, specificity=0.68). AUC value of the model was 0.69.

Later, random forest was used for diagnostic classification using 'randomForest' package in RStudio. Splitting train and test sets technique was the same as mentioned above. Number of trees should be selected for random forest, which is an important parameter that changes the accuracy level (Probst & Boulesteix, 2018). Number of trees are also selected according to trial-

and-error and the number used was 500. The maximum overall accuracy level achieved by random forest was 67%, mathematically lower than support vector machine. 66% of the patients were correctly classified, this value was 68% for HC group. (sensitivity=0.66, specificity=0.68, AUC=0.67). One important feature of random forests is Gini importance score which are calculated for each feature showing its discriminative power (Venkataraman et al., 2012). Using 'RandomForest' package on RStudio, it is possible to get a variable importance plot which computes mean decrease accuracy (MDA) and mean decrease Gini (MDG). MDA is the change in prediction accuracy for each variable when they are replaced by randomly permuted values, whereas MDG shows the decreased Gini impurity for each variable(Calle & Urrea, 2011). The MDA and MDG plots can be seen at figure 5.

Figure 5

Mean decrease accuracy and mean decrease Gini of features

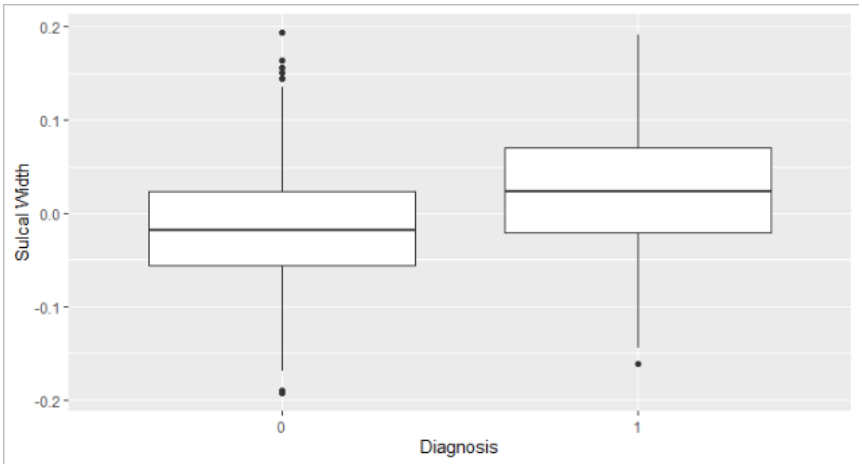


Note. Abbreviations: CT, cortical thickness; IC, intracranial volume; SI sulcation index; GM, grey matter; LGI, local gyrification index; WM, white matter.

Sulcal width turned out to be the most discriminative and important variable differentiating patients and healthy controls. Due to this finding, additional analysis has been made. Two sample t-test used to compare two groups on their sulcal width. Sulcal width values, as with any other brain imaging variables used in this study, was corrected for age and biological sex. Mean values of sulcal width differed significantly with diagnosis ($t(749.39) = 9.965$, $p < .001$). Figure 6 shows the boxplot to illustrate sulcal width values of participants, diagnosis "1" refers to the SZ group, whereas diagnosis "0" refers to the HC group. Patients, in general, had larger width. Interestingly, no relationship between sulcal width and illness severity was found using simple linear regression ($AdjR^2$=-0.002593, $F(1,351)= 0.08969$, $p=0.76$), also no relationship was found between sulcal width and illness duration of patients (AdjR2=-0.001672, $F(1,297)= 0.5024$, $p=0.47$).

Figure 6

Boxplot of sulcal width values



Note. Diagnosis "0" refers to healthy control group, whereas diagnosis "1" refers to the patient group.

Another diagnostic classification is made using logistic regression. One of the reasons of this is that machine learning methods can be seen as "black box" and easy interpretability of

logistic regression makes it an important machine learning technique (Phillips et al., 2015). The concept of black box will be discussed more extensively in the discussion section of the study. Glm function of 'stats' package was used for binomial logistic regression analysis. A table is given as a result of logistic regression indicating predictive power of each variable (Phillips et al., 2015). Out of 15 brain imaging variables, 6 of them could significantly predict the classification outcome. The variables, their coefficients and their p values are reported in table 4. In logistic regression, coefficients refer to the expected change in log odds per unit change in that variable, 1 unit increase in the value of that variable multiplies the odds of having the outcome by e^{β} (Park, 2013). Sulcal width was founded to be the most significantly associated variable, a similar result that was found with random forests. The accuracy of logistic regression was founded to be 66% (sensitivity=0.66, specificity=0.67, AUC=0.66).

Table 4

Brain imaging variables that can significantly predict diagnosis according to logistic regression analysis with their p values and coefficients.

Brain imaging variable	Coefficient (β)	P value
Cortex volume	2.6225614	0.01182
IC volume	0.5169790	0.01503
Sulcal width	0.7990807	<0.001
Sulcal length	1.0932032	0.00595
Sulcal area	-2.5133281	0.01194
Cortex SI	1.0797921	0.02240

Note. IC, intracranial volume; SI sulcation index

3.2.1 Averaging probabilities and majority voting

So far, 3 models were used for diagnostic classification: SVM, random forest and logistic regression. Their accuracy levels are 69%, 67% and 66% respectively. Ensemble methods are learning algorithms that does predictions according to votes of multiple classifiers (Dietterich, 2000). A similar method was used in this study. When making a decision, all of the mentioned methods calculate probabilities for class involvement. After training with training set, all these methods detect some patterns that differentiate schizophrenia patients from controls. When making a prediction, these models investigate the brain imaging data of each person and calculate the probability of this person belonging to each class. A new dataset was created. The probability of being a patient for each observation in the test sets were extracted from the models. Model's classification decisions were also included in the dataset. For example, if the model found a participant's probability of having schizophrenia is 80%, that person classified as diagnosed, if the probability is small, such as 20%, that person will be classified as not having schizophrenia. In the created dataset, all of these probabilities by three different methods were put together and then a new variable called "average probability" is created. The logic behind is very simple, for each participant in the test set, found probabilities by three models were averaged. To illustrate imagine a participant A. Let's say SVM, RF and LR found participant A's probability of having schizophrenia 70%, 75% and 80%. Average probability of having schizophrenia for this participant is (70+75+80)/3 =75. This calculation was made for each observation. Every observation had an averaged probability. Later, data was classified according to these probabilities. Default option of many models has a cutoff point of .5. Which means if the probability is more than 50%, the participant will be classified as diagnosed, if less than 50%, the participant is classified as undiagnosed. However, change can be made in the cutoff score for better performance (for a review, see Calle et al., 2011). Data used in the current study is

imbalanced for classes, majority of the participants were healthy controls. In such cases, smaller cutoffs than 0.5 can be used so that doubtful instances can be labelled as minority and accuracy can be improved (Zhou & Wang, 2012). Cutoff of .40 was selected as it gave the best results according to a ROC. To summarize, probability averages was calculated according to three models used, then observations which have more than 40% probability of classified as diagnosis group were labelled as diagnosed group. This gave an accuracy level of 67% (sensitivity=0.6, specificity=0.71, AUC=0.65).

Another technique, majority voting was also used. In this technique, a decision is made using the votes of the models. A participant is labelled as diagnosed only if 2 out of 3 models labelled them as patients. Accuracy level reached with this technique was 70%, (sensitivity=0.6, specificity=0.75, AUC=0.67).

Lastly, a powerful method called XGBoost was used for diagnostic classification. 'Xgboost' package in Rstudio was used for this purpose. There are more than 10 parameters of XGBOOST, however, only 2 of those parameters were controlled, and others used in the default option. These two parameters are eta and gamma. 71% accuracy was reached with XGBOOST (eta=0.16, gamma=0.06), which was the highest accuracy level reached for diagnostic classification. Accuracy of detecting healthy controls was 70% and accuracy of detecting patients were %72 (sensitivity=0.72, specificity=0.7, AUC= 0.71).

Using 'pROC' library, it is possible to calculate the minimum AUC that is significantly different than an AUC level of 0.5. When the desired p value and the number of the groups is stated, minimum AUC value to reach such significance is calculated. For diagnostic classification, this level was found to be 0.56 for p value of 0.01. According to this, AUC levels of each classification were significant.

3.3 Classification for 3 Subclinical Patterns

PANSS scale has 3 subscales called positive, negative, and general psychopathology (Kay et al., 1987). In this part of the analysis, it was investigated that if an accurate classification according to symptom patterns is possible with machine learning using brain imaging variables. Patients' data were divided into three groups:

1) Predominant general psychopathology symptoms group

2) Predominant negative symptoms group

3) Predominant positive symptoms group

This division was made according to which symptom pattern they suffer the most according to their PANSS scores. Their score for each subscale is first calculated, later, these scores were scaled since general psychopathology subscale has more items. The maximum score for GP subscale is 112, whereas for positive and negative subscale the maximum score is 49. Their score for each subscale is divided by the maximum score of that subscale. Later, data of patients were assigned one of these three groups according to from which subscale they had the maximum score. This dataset contained 350 observations; 41 of them were assigned as GP symptoms group, 160 of them assigned as negative symptoms group and 149 of them assigned as positive symptom group. Later, dataset was divided as train and test set in a way that no brain imaging from the same person will end up being in different sets.

First, a radial kernel support vector machine algorithm was used for classification, parameters of support vector machine was found with trial and error and best-balanced accuracy was caught by cost=1 and gamma=0.4. Overall accuracy was 48% (Using the package pROC, it is possible to get AUC values even for multiclass classification. AUC for this analysis was 0.69. Algorithm was most successful detecting brain imaging data that belongs to patients with predominant

negative symptoms, this followed by positive and general psychopathology respectively. This could be due to smaller data in general psychopathology group. Test set of GP group included brain imaging variables of only 6 people, 2 of whom was correctly detected by the algorithm. Algorithm could accurately classify 55% of negative group and it could accurately classify 53% of positive group.

Random forest was also used for this classification since SVM is generally avoided when dealing with more than 2 classes (Duan & Keerthi, 2005). After implementation of same train-test set division technique, random forests with 500 trees yielded a maximum overall accuracy of 55%. Random forest was mathematically more successful than SVM on overall accuracy. Furthermore, 71% of people in the negative group were correctly classified, and 50% of people in positive group were correctly classified. However, none of the brain imaging data that belongs to GP group was correctly classified. Interestingly, both models were more successful at the classification of brain imaging data of negative group. AUC of the model was 0.58. According to variable importance plots of random forests, sulcal width was again the most important variable for differentiation for the model, however the mean differences between groups was not significant as shown by ANOVA ($F(2,347) = 1.109$, $p=0.331$).

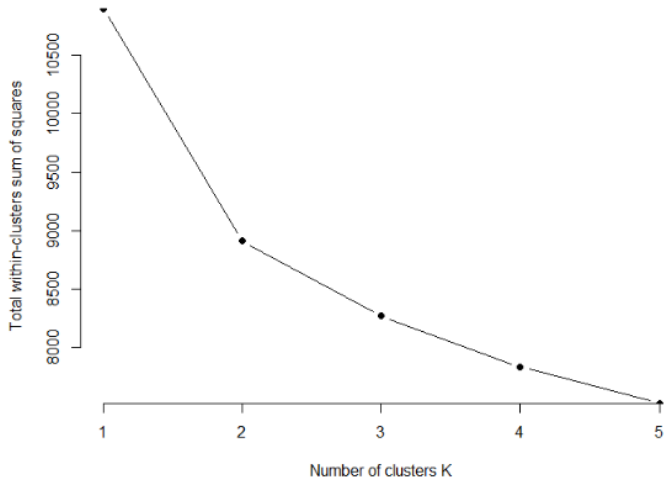
3.4 Unsupervised Model for Subgrouping Patients

Some researchers found different number of factors for PANSS scale factors (Emsley et al., 2003; Van den Oord et al., 2006; Kim et al., 2012, Levine, & Rabonowitz, 2007; Wallwork et al., 2012; Chen et al., 2020). In order to see if there are different numbers of symptom patterns for patients in PANSS and to see if a new division can achieve a better classification, various models were used in this study. First of all, for clustering, the number of classes should be

determined before the analysis, and one possible way of doing that is elbow method using k-means clustering (Andrew, 2012). It is a visual model for deciding the number of clusters, after starting with k=2 and keep increasing the number of clusters, the cost should be monitored. The k value is selected at the point where there is a dramatic drop of cost (Kodinariya & Makvana, 2013). 'ClusterR' package was used for this purpose. First PANSS scores of patients were used for k-means clustering, later, to employ the elbow method, this process was visualized. The number of clusters and their costs was shown in figure 7. According to this visualization, and the direction were given in Kodinariya and Makvana (2013), 2 clusters were selected as k. This means, according to PANSS scores of patients, there are 2 subclasses.

Figure 7

The number of clusters and their costs



For further investigation, a factor analysis with 2 factors using 'Stats' library in RStudio was conducted. The resulting table shows the magnitude of relationship of each item with each factor. As stated by Pett et al., (2003) labelling of factors is a process that is subjective and theoretical (as cited in Williams et al., 2010). Thus theoretical knowledge was used for labelling of factors. First, Kay et al., (1987), while proposing their 3 factors scale, stated that general

psychopathology subscale was not meant to be a distinct subscale, but it is an important addition of positive/negative subscales. Additionally, positive/negative symptom dichotomy is commonly mentioned for schizophrenia (e.g., Crow, 1980; Lieberman et al., 2020; APA, 2013). Furthermore, the results of factor analysis show each item's relationship with each factor and investigation of the table showed positive/negative dichotomy between factors. More precisely, every item that originally belongs to positive subscale of PANSS showed stronger relationship with one factor, whereas nearly all items that originally belongs to negative subscale showed stronger relationship with the other factor. Considering these, emerging factors were decided to show the positive/negative dichotomy of schizophrenia.

According to their loadings (magnitude of correlations with each factor), each item was assigned to either negative or positive factor. The results were shown in table 5. Only 1 item, stereotyped thinking, that originally belongs to negative subscale included in the positive factor due to its stronger relationship with positive symptom pattern. As can be seen, more items were correlated with positive symptom pattern.

Table 5

Items and their assigned factors

Factor 1- Negative	Factor 2- Positive
Blunted affect	Delusions
Emotional withdrawal	Conceptual disorganization
Poor rapport	Hallucinations
Passive apathetic social withdrawal	Excitement
Lack of spontaneity	Grandiosity
Mannerism and posturing	Suspiciousness persecution

Depression	Hostility
Motor retardation	Stereotyped thinking
Disturbance of volition	Somatic concern
Disorientation	Guilt feelings
Active social avoidance	Tension
Difficulty abstract thinking	Anxiety
	Unusual thought content
	Poor impulse control
	Uncooperativeness
	Poor attention
	Lack of judgement and insight
	Preoccupation

Cronbach's alpha of PANSS scale was 0.884, which was found using 'ltm' library (Rizopoulos, 2007). Confirmatory factor analysis (CFA) can be conducted to see the fit of the factor models (Sarmento, & Costa, 2019). Therefore, to see the difference between 2-factor and 3-factor model, confirmatory factor analysis was conducted. CFA was conducted using 'lavaan' library in RStudio (Rosseel, 2012). Fit indices showed that comparing two-factor model (X2 = 1635.107, df. = 404; RMSEA = 0.100; CFI = 0.61) and three-factor model (X2= 1778.724, df= 402, RMSEA=0.106; CFI=0.57) is a little bit complicated. Sarmento and Costa (2019) stated X^2 (Chi-squared test) value indicates the difference between expected and observed covariance matrices, a smaller value is better. Similarly, smaller values indicate success with Root Mean Square Error of Approximation (RMSEA). CFI and RMSEA, evaluates the model adjusting for

sample size, as it can be an issue for Chi-squared test. Higher CFI refers to a better fit. In terms of RMSEA and CFI, two factor model is a better fit, however, X^2 value of three factor model is better. It is important to note here that both models failed to be considered as a good model. More information about appropriate values for X^2, CFI and RMSEA can be found in (Sarmento, & Costa, 2019).

Later patient's data were divided into 2 according to from which symptom pattern they show higher scores. Train and test set division was conducted in the same way as for all classifications. A support vector machine with radial kernel with cost=0.01 and gamma=0.4, yielded the best results. The overall accuracy was 61%; 66% of brain imaging variables were correctly classified as belonging to the negative group and 55% of brain imaging variables were correctly classified as belonging to the positive group. (AUC=0.61)

To be in line with the previous analysis, a random forests classification with 1000 trees was conducted. Overall accuracy was 66%; 75% of data belonging to negative group was correctly classified and 55% of the data belonging to positive group was correctly classified. AUC of the model was 0.65. Two observations can be made with these results. Mathematically speaking, two symptom patterns increased accuracy for classification. Secondly, detection power of negative group overcame the positive group, similar to the previous classification. According to random forest variable importance plot, sulcal depth was the most important variable differentiating two groups. Mean of sulcal depth was higher for people with predominantly positive symptoms but t-test showed that the difference between two groups was not significant ($t(285.34) = -1.696$, $p=0.09$).

Again, minimum required AUC levels were calculated. To achieve a significance level of 0.01, minimum AUC should be 0.71 for 3-group classification and it should be 0.63 for 2-group

classification. Only AUC value of RF for 2-class classification exceeds the minimum AUC level for significance. This partly shows that 2-class classification outperformed 3-class classification.

3.5 Classification according to illness duration

Illness duration was divided into two for methodological reasons. The success of machine learning algorithms drops when handling with unbalanced datasets (Ganganwar, 2012). Furthermore, the importance of sample sizes was mentioned many times throughout the paper. 296 brain imaging observations were available for illness duration, which is a large dataset in general compared to other studies done in the field. However, for machine learning classification, dataset should be divided as train and test sets and in this study 20% of the dataset was used as a test set and it included 52 observations, which would be smaller if illness duration was divided into more than two groups. Thus, brain imaging data was divided into two, according to illness durations. The median of illness duration was 7.79 years so division was made at 7.79 years. Data of people with an illness duration less than 7.79 years were labelled as "shorter illness duration group" and data of people with an illness duration more than 7.79 years were labelled as "longer illness duration group". Support vector machine with radial kernel was trained with brain imaging variables of people with shorter and longer illness duration and later, the algorithm was tested with a test set. According to brain imaging variables, it predicted if that brain imaging values belongs to a person with shorter illness duration or longer illness duration. The overall accuracy for this classification was 75%. 60% of the data in longer illness duration group was correctly classified. 86% of the data in shorter illness duration group was correctly classified. AUC value for this classification was 0.73. Minimum AUC level for a significance level of 0.01 was found to be 0.63, which is less than the AUC level of the classifier.

3.6 Classification According to Illness Severity

PANSS scores were used as an indication for illness severity of patients (Kay et al., 1987). Due to the methodological concerns stated at previous classification for illness duration, illness severity was also divided into two. Median PANSS score was 54, and patients who scored less than 54 were labelled as less severe SZ group and patients who scored more than 54 were labelled as more severe schizophrenia group. Support vector machine with radial kernel was used for classification and this classification yielded an accuracy level of 69%. Cost and gamma levels were 0.001 and 0.15 respectively. AUC value for this classification was 0.69. %65 of data in less illness severity group was correctly classified, and %72 of the data in more severe group was correctly classified. Minimum AUC level for a significance level of 0.01 was found to be 0.62, which is less than the AUC level of the classifier.

For better visualization, each classification and their results are displayed on table 6. results will be discussed in the next section of the paper.

Table 6

Used methods, their accuracy and AUC levels.

Algorithm	Classification type	Overall accuracy (95% CI)	Controlled parameters	AUC	Sample size (Number of brain imaging variables)

SVM	Diagnosis	69.2% (62-75%)	Gamma=0.008, cost=0.006	0.69	1087
RF	Diagnosis	67% (61-74%)	nTrees=500	0.67	1087
Logistic R.	Diagnosis	66% (60-73%)	-	0.66	1087
XGBoost	Diagnosis	71% (64-77%)	eta=0.16, gamma=0.06	0.71	1087
SVM	Symptom pattern	48% (36-60%)	cost = 1, gamma = 0.4	0.69	350
RF	Symptom pattern	55% (43-67%)	nTrees=500	0.58	350
SVM	Symptom pattern*	61% (48-73%)	cost=0.01, gamma=0.4	0.61	302
RF	Symptom pattern*	66% (53-78%)	nTrees=1000	0.65	302
SVM	Illness duration	75% (61-85%)	cost=0.4, gamma=0.09	0.73	296
SVM	Illness severity	69% (56-80%)	cost=0.001, gamma=0.15	0.69	353

*Note. *symptom pattern for 2-class classification.*

CHAPTER 4

DISCUSSION

Schizophrenia is a chronic and disabling disorder (Mueser & Jeste, 2008) with 13.5% recovery rate (Jääskeläinen et al., 2013). Worst outcomes can be prevented through early detection and early intervention (Razzouk et al., 2006), however, due to the absence of an objective diagnostic tool, this is currently not possible (Bell et al., 1998). Furthermore, Wolfers and colleagues (2018) suggests the treatment failure could be due to one average patient group assumption for a group that shows high heterogeneity. Individualized medicine got attention during the last years due to awareness of individual differences between patients and the importance of finding an optimal treatment plan for each individual (Wang, 2012). Predictions on individual level is possible through machine learning (Bzdok & Meyer-Lindenberg, 2018). Thus, this study used various machine learning models to make predictions about diagnosis of

schizophrenia, subtypes, illness duration and illness severity using structural brain imaging data of the participants.

Four different models were used for diagnostic classification. These methods were support vector machines, random forests, logistic regression, and XGBoost. Each model was first trained with brain imaging data of participants so models could detect some patterns. Later a part of brain imaging dataset was used for testing, in which model made a prediction for each observation about being belong to a patient or not, accordingly to the patterns found in training. For diagnostic classification, XGBoost had the best overall accuracy with 71%. This was followed by a majority voting technique. In this technique, all SVM, RF and logistic regression were used to make a decision. Brain imaging data of a participant were classified as schizophrenia group only when two out of three models classified this participant as SZ group. The accuracy level reached by this method was 70% with AUC=0.67. This method outperformed all three methods in terms of accuracy; however, AUC value is lower than SVM, equal to RF and higher than logistic regression. In between those three models, SVM performed the best by 69.2% overall accuracy and logistic regression performed worst with 66% overall accuracy. Analysis with neuroimaging data is generally challenging due to high dimensional nature of such data and small sample sizes (Menoret et al., 2018). Although the sample size of the current study is larger compared to many similar studies, classification studies tend to use larger datasets in other fields. An example is a classification study for spam e-mail, and they had a sample size of 98680 (Chen et al., 2019). Support vector machines shown to be powerful when working with smaller and high-dimensional datasets (Yan et al., 2004), Thus, better accuracy outcome by support vector machine could be explained in this way. Furthermore, SVM, is less time consuming and it is a simpler method compared to, for example, XGBoost. Due to simplicity and

higher accuracy level reached, SVM models can be suitable to use as diagnostic tools for schizophrenia. Majority voting technique used in the study reached an accuracy of 70%, which was not a great improvement considering 69% of accuracy level was already reached by SVM. According to the results of these models, hypothesis 1, supervised machine learning methods can accurately predict if people have schizophrenia using their structural brain imaging variables, was accepted.

According to variable importance plots, sulcal width was a very important measure to differentiate SZ and HC group. This was confirmed by a t-test, where the means of sulcal width of patients and healthy controls were shown to be significantly different. Furthermore, logistic regression analysis showed that out of fifteen brain imaging variables, six of them could significantly predict the classification outcome. Sulcal width was the variable which had the smallest p value, meaning it was the most important variable for this differentiation. Thus, these results suggest, sulcal width could be a biomarker for schizophrenia. This finding was also reported by Janssen and colleagues (2014), they stated an increased sulcal width was seen both in patients with schizophrenia and bipolar disorder. They stated a connection between decreased gyrification, increased sulcal width and schizophrenia. Increased sulcal width was found to be correlated with advanced age (Kochunov et al., 2005), and Alzheimer's disease (Hamelin et al., 2015). It is also found to be associated with worse cognitive function in a healthy population (Liu et al., 2011). These findings including the results of the current study suggest the importance of sulcal width.

Machine learning methods were also used to classify patients according to their symptom patterns. PANSS scale originally have three symptom patterns that are positive, negative and general psychopathology. First a classification was made using these three original symptom

patterns. Brain imaging data of participant were divided into three according to their maximum score on these subscales. Later, SVM and RF models were trained and tested to classify patients' data for symptom patterns. The accuracy of SVM was 48% with 0.69 AUC. RF reached an accuracy of 55%, however, the AUC value of the model was worse than SVM, 0.58. Later, K-means clustering was used to identify number of factors in PANSS scale and the number of factors was two. A factor analysis took place to find each factors and related items. Factor analysis yielded positive/negative dichotomy, which is commonly mentioned for schizophrenia. Thus, according to their PANSS scores, patients were divided into two groups: patients with predominantly positive symptoms and patients with predominantly negative symptoms. For comparison reasons, just like with the other subgroup classification, SVM and RF were used for classification. SVM reached an accuracy of 61% with 0.61 AUC and RF reached 66% of accuracy with 0.65 AUC. For the first analysis with 3 subgroups, sulcal width was again the most important variable for differentiation. For the latter analysis with 2 subgroups, sulcal depth was the most important variable, however, the group differences were not significant according to t-test and ANOVA. The related hypotheses about this section were:

Hypothesis 2: Supervised machine learning methods can accurately classify patients into three groups according to their symptom patterns using their structural brain imaging data.

Hypothesis 3: Classification of patients according to their symptom patterns through an unsupervised machine learning method can increase the accuracy of MRI-based subtype classification.

Here, the evaluation of the hypothesis should be done carefully. First of all, accuracy levels were higher for 2-group classifications than 3-group classifications, however, AUC of SVM model for 3-group classification outperformed every other classification. Another way to

comment on this issue could be finding a minimum AUC for significance. In 'pROC' library, it is possible to calculate the minimum AUC to be significantly different than an AUC level of 0.5. To achieve this calculation, number of observations in groups and the desired p value should be stated. According to this calculation, to achieve a significance level of 0.01, minimum AUC should be 0.71 for 3-group classification and it should be 0.63 for 2-group classification. Only AUC value of RF for 2-class classification exceeds the minimum AUC level for significance. Thus, according to this result, hypothesis 2 should be rejected and hypothesis 3 should be accepted. Unsupervised machine learning method can increase the accuracy of MRI-based subtype classification.

Two SVM models were created for classification for illness severity and illness duration. Patients were divided into two according to their illness severity and illness duration from their median values for methodological reasons. First reason is that sample size would be smaller with higher number of groups, and this could be problematic for testing part. Secondly, unbalanced dataset may result in problematic accuracy level in machine learning (Ganganwar, 2012). By dividing patients into two using the median levels for illness duration and illness severity, it was possible to create balanced group sizes with larger sample size. SVM with radial was used for both classifications and their accuracy levels were 69% and 75% for illness severity and illness duration respectively. Their AUC values were 0.69 and 0.73 for illness severity and illness duration respectively. Both hypothesis 4 and 5 were accepted because it was possible to classify patients according to their illness duration and illness severity.

4.1 Significance of the Current Study

Throughout the paper, the importance of an objective diagnostic tool for schizophrenia is mentioned. The main aim of the study was to create an objective tool for diagnosis. Accuracy

levels for diagnostic classification differed between 66% to 71%. Winterburn and colleagues (2019), in their paper, investigated 18 machine learning diagnostic classification studies that used MRI. The accuracy levels differed between 63.2% to 93.3%. The reached accuracy levels in this study are generally lower than the mentioned accuracy levels in these studies. However, sample sizes used for classification were generally smaller than the current study. Mean sample size of 18 studies was 118, whereas current study had a sample size of 467, and since image acquisition took place more than one time for some participants, 1087 brain images included in the study. Schnack and Kahn (2016) stated small sample sizes tend to result in a classification with higher accuracy and low generalizability and vice versa for larger samples. Since the aim of the study is to create a diagnostic tool for clinical use, generalizability is really important. Thus, this is a crucial study, using a large sample for classification, which may result in better generalizability. However, machine learning application in mental health have a relatively short history and the algorithms are not yet ready to be used in clinics. Salgado-Pineda and colleagues (2018) stated, AUC levels between 0.6 and 0.7 can be considered as poor, AUC levels between 0.7 and 0.8 can be considered fair. According to this statement, most of the diagnostic classification in this study should be considered poor. AUC levels for diagnostic classification differed between 0.66 to 0.71. An algorithm with a fair classification is not ready to be used in the clinics, because misclassification would be costly. However, as mentioned, machine learning applications are relatively new in mental health, so we are getting there. This study has a significant importance for using a large sample, which makes the tools to have higher generalizability power. Moreover, in this study four different methods were used for the same sample, which gives clues about which method could work better for schizophrenia for later use in clinics.

In this study, participants were also classified according to their symptom patterns. The importance of finding subgroups of patients were stated numerous times in the paper. Every patient with schizophrenia does not show the same symptoms, they have many differences and the underlying abnormalities in the brain may differ. A discovery of stable subgroups can be a huge help. In this study, it has been showed that it was possible to detect patients with predominantly positive symptoms vs patients with predominantly negative symptoms. This tells us these groups differ in terms of their structural brain patterns and this difference, through further research, could lead to more accurate and appropriate intervention. The needs of these different groups can be found through research, and machine learning tools may help clinicians to discover the subgroup of the patient. Antipsychotics were shown to be effective with positive symptoms but not effective with negative symptoms (Mueser et al., 2013). It is clear that the needs of patients who suffer mostly from negative symptoms differ than the patients who suffer mostly from positive symptoms. The current study confirms there are different structural brain imaging patterns that are detectable for these different groups. Chen et al. (2020) also stated similar results. They found four factors for schizophrenia symptoms which were positive, negative, affective, and cognitive dimensions. However, positive, and negative psychopathological subtypes were found to be stable and core subtypes, and it could be discriminated by resting state functional connectivity patterns of patients. Model reached an accuracy of 70%.

Classification for illness duration and illness severity were not performed for clinical use aim. Although there is still room for error due to late diagnosis, generally illness duration is known by the patient and the clinician. Therefore, a machine learning algorithm which detects the illness duration would not be useful in clinics. However, the ability of the machine to

correctly classify people according to their illness duration using their structural brain images confirms that illness duration changes structural brain patterns of the patients. As stated, before the analysis, brain imaging variables were corrected for participants' age. Thus, the different patterns found can not be explained with age differences. Similarly, a machine learning algorithm to detect illness severity may not be useful in the clinics as it is known. However, since this study showed it is possible, with further research, such classification could be used in clinics as well in the future. A longitudinal dataset with large sample size is needed for this aim. With such dataset, it could be possible to predict future illness severity of patient and means of intervention could be increased.

4.2 Limitations and Future Studies

There are some limitations of the current study. First, although they are powerful methods, there are some limitations of using machine learning. The aim of many studies which use machine learning, including this study, is to find a general predictive rule through the algorithms and use them later to make predictions. However, some machine learning models can work well with the dataset they are trained with, but the success cannot be seen in another dataset. This is called overfitting (Dietterich, 1995). The parameters of machine learning models were found using a trial-and-error approach. Different parameters for each model were used until the maximum accuracy is reached on the test set. However, this does not guarantee that these selected parameters would show the same performance on another dataset. These parameters could be the best parameters for the test set that was used in the current study, but they may not be the best in general for detecting schizophrenia. Luckily, usage of large sample decreases this risk (Schnack, 2017).

Another highly discussed limitation of machine learning is that they are seen as “black box” (see figure 8). The decision-making process is hidden in machine learning algorithms, so there is an accurate model which can correctly classify groups, but we do not know how (Poulet, 2005). All of the methods used for classification in this study can be considered as black box (Ljubobratović & Matetić, 2019; LaValley, 2008). However, especially in medical research, explanations about the data is very important. We want to know if it is possible to use a tool for classification, but we also want to know how the algorithm could make such differentiation as well to explore patterns related to the disorders. Unfortunately, this is not possible with machine learning methods alone. Some efforts have been shown to make these models more interpretable (see Poulet, 2005). However, this is not the case for the current study. Only, for diagnostic classifications, logistic regression was also chosen to be one of the methods because they are more interpretable (Kung & Yu, 2020). Although this can change with the usage of different libraries and different programming languages, with the current used libraries, random forest and logistic regression were found to be the most interpretable ones. Both models state the importance of the variables easily and logistic regression even show the significance levels of each variable for classification. SVM was used more often due to higher achieved accuracy and simplicity of the model, however, usage of random forests and logistic regression for every classification could make the current research more interpretable and more informative. Future studies can use multiple methods for each classification to reach more interpretable results.

Figure 8

Parameter selection was conducted with a trial-and-error approach. The default values from the libraries were used as a starting point and different values were tried until the best accuracy is achieved using the test set. Although this is a used approach, there are more sophisticated approaches for parameter selection (see Bergsta, & Bengio, 2012). Future studies can use tuning functions for parameter selection for better results. Tuning is possible using 'caret' library in R programming language.

Cross validation is a commonly used technique which splits the data as train and test set differently numerous time to better understand the predictive power of the model and to avoid overfitting (Arlot & Celisse, 2010). Cross validation could not be performed due to the complexity of data. In the dataset, there are some participants who had their image acquired two or three times. As explained in the method section, train and test sets were divided in a way that no participant's brain images would end up in both sets. Creating the same situation for cross validation was very complex. Using cross validation without controlling this concept would create error because the brain imaging values of the same participant could end up in both train and test sets, which would make these sets less independent. Not using an independent test set would lead to an unfair comparison (Kelly et al., 2019).

Finally, comments on significance of the classifier were done using minimum AUC level for significance and comparing them mathematically to the AUC level of the classifier, however, there are more sophisticated methods for such testing. Permutation testing is a technique that can be used for machine learning methods where the performance of the classifier is compared with numerous random (permuted) classifications (Westerhuis et al., 2008). This way, it is possible to understand if classifier can do a better job than randomly assigning classes. Permutation testing

was not performed due to the complexity and computational cost. Future studies can use permutation testing to have a more sophisticated evaluation.

4.3 Conclusion

The aim of this master thesis was to create objective tools using machine learning for schizophrenia. Methods such as support vector machines, random forests, logistic regressions, and XGBoost were used make predictions about diagnosis, disease subtypes, illness duration, and illness severity using their MRI data. Machine learning was chosen due to it is applicability in real life. It is hoped that this research will help people with schizophrenia in the future.

References

Ab Kadera, N. I., Yusofa, U. K., & Naima, S. (2019). Diabetic retinopathy classification using support vector machine with hyperparameter optimization. *Int. J. Advance Soft Compu. Appl*, *11*(3).

Adityanjee, Aderibigbe, Y. A., Theodoridis, D., & Vieweg, W. V. (1999). Dementia praecox to schizophrenia: The first 100 Years. *Psychiatry and Clinical Neurosciences*, *53*(4), 437–448. https://doi.org/10.1046/j.1440-1819.1999.00584.x

Adriano, F., Spoletini, I., Caltagirone, C., & Spalletta, G. (2010). Updated meta-analyses reveal thalamus volume reduction in patients with first-episode and chronic schizophrenia. In *Schizophrenia Research*, 123(1), 1–14. https://doi.org/10.1016/j.schres.2010.07.007

Adriano, F., Caltagirone, C., & Spalletta, G. (2012). Hippocampal volume reduction in first-episode and chronic schizophrenia: A review and meta-analysis. *Neuroscientist* 18(2), 180–200. https://doi.org/10.1177/1073858410395147

Ahmed, A. O., Strauss, G. P., Buchanan, R. W., Kirkpatrick, B., & Carpenter, W. T. (2018). Schizophrenia heterogeneity revisited: Clinical, cognitive, and psychosocial correlates of statistically-derived negative symptoms subgroups. *Journal of Psychiatric Research*, *97*, 8–15. https://doi.org/10.1016/j.jpsychires.2017.11.004

Al-Mejibli, I. S., Alwan, J. K., & Abd Dhafar, H. (2020). The effect of gamma value on support vector machine performance with different kernels. *International Journal of Electrical and Computer Engineering*, *10*(5), 5497. doi: 10.11591/ijece.v10i5.pp5497-5506

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). https://doiorg.ezproxy.frederick.edu/10.1176/appi.books.9780890425596

Andrew, N.g., (2012). Clustering with the K-Means Algorithm, Machine Learning,

Arlot, S., & Celisse, A. (2010). A survey of cross-validation procedures for model selection. *Statistics surveys,* 4, 40-79. doi: 10.1214/09-SS054

Assaf, M., Rivkin, P. R., Kuzu, C. H., Calhoun, V. D., Kraut, M. A., Groth, K. M., ... & Pearlson, G. D. (2006). Abnormal object recall and anterior cingulate overactivation correlate with formal thought disorder in schizophrenia. *Biological psychiatry*, 59(5), 452-459. https://doi.org/10.1016/j.biopsych.2005.07.039

Ayano, G., Tesfaw, G., & Shumet, S. (2019). The prevalence of schizophrenia and other psychotic disorders among homeless people: A systematic review and meta-analysis. *BMC Psychiatry*.19(1). BioMed Central Ltd. https://doi.org/10.1186/s12888-019-2361-7

Baiano, M., David, A., Versace, A., Churchill, R., Balestrieri, M., & Brambilla, P. (2007). Anterior cingulate volumes in schizophrenia: A systematic review and a meta-analysis of MRI studies. *Schizophrenia Research*, *93*(1–3), 1–12. https://doi.org/10.1016/j.schres.2007.02.012

Baldessarini, R. J., & Frankenburg, F. R. (1991). Clozapine: A novel antipsychotic agent. *The New England Journal of Medicine, 324*(11), 746–754. https://doi.org/10.1056/NEJM199103143241107

Bansal, R., Staib, L. H., Laine, A. F., Hao, X., Xu, D., Liu, J., Weissman, M., Peterson, B. S., & Zhan, W. (2012). Anatomical Brain Images Alone Can Accurately Diagnose Chronic Neuropsychiatric Illnesses. *PLoS ONE*, *7*(12). https://doi.org/10.1371/journal.pone.0050698

Bansal, V., & Chatterjee, I. (2021). Role of neurotransmitters in schizophrenia: A comprehensive study. Kuwait Journal of Science, 48(2). https://doi.org/10.48129/kjs.v48i2.9264

Barrera, A., Mckenna, P. J., & Berrios, G. E. (2005). Formal thought disorder in schizophrenia: an executive or a semantic deficit?. *Psychological Medicine*, 35(1), 121-132. https://doi.org/10.1017/S003329170400279X

Barrett, L. F. (2018). Seeing fear: It's all in the eyes? *Trends in Neurosciences*, 41(9), 559–563. https://doi.org/10.1016/j.tins.2018.06.009

Bauer, S. M., Schanda, H., Karakula, H., Olajossy-Hilkesberger, L., Rudaleviciene, P., Okribelashvili, N., ... & Stompe, T. (2011). Culture and the prevalence of hallucinations in schizophrenia. *Comprehensive psychiatry*, *52*(3), 319-325. https://doi.org/10.1016/j.comppsych.2010.06.008

Becker, T. M., Cicero, D. C., Cowan, N., & Kerns, J. G. (2012). Cognitive control components and speech symptoms in people with schizophrenia. *Psychiatry Research*, *196*(1), 20–26. https://doi.org/10.1016/j.psychres.2011.10.003

Behrendt, R. P., & Young, C. (2004). Hallucinations in schizophrenia, sensory impairment, and brain disease: A unifying model. *Behavioral and Brain Sciences*, *27*(6), 771–787. https://doi.org/10.1017/S0140525X04000184

Bell, R. C., Dudgeon, P., McGorry, P. D., &; Jackson, H. J. (1998). The dimensionality of schizophrenia concepts in first-episode psychosis. *Acta Psychiatrica Scandinavica*, 97(5), 334–342. https://doi.org/10.1111/j.1600-0447.1998.tb10012.x

Bell, M. D., Johannesen, J. K., Greig, T. C., & Wexler, B. E. (2010). Memory profiles in schizophrenia: Categorization validity and stability. *Schizophrenia research*, *118*(1-3), 26-33. https://doi.org/10.1016/j.schres.2009.12.037

Benba, A., Jilbab, A., & Hammouch, A. (2017). Using Human Factor Cepstral Coefficient on Multiple Types of Voice Recordings for Detecting Patients with Parkinson's Disease. IRBM, 38(6), 346–351. https://doi.org/10.1016/j.irbm.2017.10.002

Ben-Hur, A., Ong, C. S., Sonnenburg, S., Schölkopf, B., & Rätsch, G. (2008). Support vector machines and kernels for computational biology. *PLoS Computational Biology*, *4*(10). https://doi.org/10.1371/journal.pcbi.1000173

Bentall, R. P., Baker, G. A., and Havers, S. (1991). Reality monitoring and psychotic hallucinations. Br. *J. Clin. Psychol*. 30, 213–222. doi: 10.1111/j.2044-8260.1991.tb00939.x

Bergstra, J., & Bengio, Y. (2012). Random search for hyper-parameter optimization. *Journal of machine learning research*, 13(2).

Bhavsar, H., & Panchal, M. H. (2012). A review on support vector machine for data classification. *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)*, *1*(10), 185-189.

Bleuler E. Dementia Praecox oder Gruppe der Schizophrenien. Deiticke, Leipzig, 1911. [Dementia Praecox or the Group of Schizophrenias. Translated by J Zinkin. International University Press, New York, 1950.]

Bourque, F., van der Ven, E., & Malla, A. (2011). A meta-analysis of the risk for psychotic disorders among first- and second-generation immigrants. *Psychological Medicine*, *41*(5), 897–910. https://doi.org/10.1017/S0033291710001406

Bornstein, R. A., Schwarzkopf, S. B., Olson, S. C., & Nasrallah, H. A. (1992). Third-ventricle enlargement and neuropsychological deficit in schizophrenia. *Biological Psychiatry*, *31*(9), 954–961. https://doi.org/10.1016/0006-3223(92)90121-f

Brébion, G., Stephan-Otto, C., Ochoa, S., Roca, M., Nieto, L., & Usall, J. (2016). Impaired Self-Monitoring of Inner Speech in Schizophrenia Patients with Verbal Hallucinations and in Non-clinical Individuals Prone to Hallucinations. *Frontiers in psychology*, *7*, 1381. https://doi.org/10.3389/fpsyg.2016.01381

Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5-32.

Briskman, I., Bar, G., Boaz, M., & Shargorodsky, M. (2012). Impact of co-morbid mental illness on the diagnosis and management of patients hospitalized for medical conditions in a general hospital. *International Journal of Psychiatry in Medicine*, *43*(4), 339–348. https://doi.org/10.2190/PM.43.4.d

Borgwardt, S., Koutsouleris, N., Aston, J., Studerus, E., Smieskova, R., Riecher-Rössler, A., & Meisenzahl, E. M. (2013). Distinguishing prodromal from first-episode psychosis using

neuroanatomical single-subject pattern recognition. *Schizophrenia bulletin*, *39*(5), 1105-1114., https://doi.org/10.1093/schbul/sbs095

Brüne, M. (2005). “Theory of mind” in schizophrenia: A review of the literature. In *Schizophrenia Bulletin*, 31(1), 21–42. https://doi.org/10.1093/schbul/sbi002

Bryson, G., Whelahan, H. A., & Bell, M. (2001). Memory and executive function impairments in deficit syndrome schizophrenia. *Psychiatry Research*, *102*(1), 29–37. https://doi.org/10.1016/s0165-1781(01)00245-1

Bzdok, D., & Meyer-Lindenberg, A. (2018). Machine learning for precision psychiatry: opportunities and challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *3*(3), 223-230. https://doi.org/10.1016/j.bpsc.2017.11.007

Cahn, W., Rais, M., Stigter, F. P., van Haren, N. E. M., Caspers, E., Hulshoff Pol, H. E., Xu, Z., Schnack, H. G., & Kahn, R. S. (2009). Psychosis and brain volume changes during the first five years of schizophrenia. *European Neuropsychopharmacology*, *19*(2), 147–151. https://doi.org/10.1016/j.euroneuro.2008.10.006

Calle, M. L., & Urrea, V. (2011). Letter to the editor: Stability of Random Forest importance measures. In *Briefings in Bioinformatics*, 12(1), 86–89. https://doi.org/10.1093/bib/bbq011

Calle, M. L., Urrea, V., Boulesteix, A. L., & Malats, N. (2011). AUC-RF: A new strategy for genomic profiling with random forest. *Human Heredity*, *72*(2), 121–132. https://doi.org/10.1159/000330778

Carpenter Jr, W. T., Heinrichs, D. W., & Alphs, L. D. (1985). Treatment of negative symptoms. *Schizophrenia bulletin*, *11*(3), 440-452.

Chance, S. A., Esiri, M. M., & Crow, T. J. (2002). Amygdala volume in schizophrenia: Post-mortem study and review of Magnetic Resonance Imaging Findings. British Journal of Psychiatry, 180(4), 331–338. https://doi.org/10.1192/bjp.180.4.331

Chang, W. C., Hui, C. L. M., Tang, J. Y. M., Wong, G. H. Y., Lam, M. M. L., Chan, S. K. W., & Chen, E. Y. H. (2011). Persistent negative symptoms in first-episode schizophrenia: A prospective three-year follow-up study. *Schizophrenia Research*, *133*(1–3), 22–28. https://doi.org/10.1016/j.schres.2011.09.006

Charlson, F. J., Ferrari, A. J., Santomauro, D. F., Diminic, S., Stockings, E., Scott, J. G., McGrath, J. J., & Whiteford, H. A. (2018). Global epidemiology and burden of schizophrenia: Findings from the global burden of disease study 2016. *Schizophrenia Bulletin*, *44*(6), 1195–1203. https://doi.org/10.1093/schbul/sby058

Chatterjee, I., Kumar, V., Rana, B., Agarwal, M., & Kumar, N. (2020). Identification of changes in grey matter volume using an evolutionary approach: an MRI study of schizophrenia. *Multimedia Systems*, *26*(4), 383–396. https://doi.org/10.1007/s00530-020-00649-6

Chaudhury, S., Kumar, S., & Soren, S. (2009). Hallucinations: Etiology and clinical implications. *Industrial Psychiatry Journal*, *18*(2), 119. https://doi.org/10.4103/0972-6748.62273

Chen, J., Patil, K. R., Weis, S., Sim, K., Nickl-Jockschat, T., Zhou, J., Aleman, A., Sommer, I. E., Liemburg, E. J., Hoffstaedter, F., Habel, U., Derntl, B., Liu, X., Fischer, J. M.,

Kogler, L., Regenbogen, C., Diwadkar, V. A., Stanley, J. A., Riedl, V., … Visser, E. (2020). Neurobiological Divergence of the Positive and Negative Schizophrenia Subtypes Identified on a New Factor Structure of Psychopathology Using Non-negative Factorization: An International Machine Learning Study. *Biological Psychiatry*, *87*(3), 282–293. https://doi.org/10.1016/j.biopsych.2019.08.031

Chen, J., Patil, K. R., Weis, S., Sim, K., Nickl-Jockschat, T., Zhou, J., ... & Visser, E. (2020). Neurobiological divergence of the positive and negative schizophrenia subtypes identified on a new factor structure of psychopathology using non-negative factorization: an international machine learning study. *Biological psychiatry*, *87*(3), 282-293. https://doi.org/10.1016/j.biopsych.2019.08.031

Chen, T., & Guestrin, C. (2016). XGBoost: A scalable tree boosting system. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, T., He, T., Benesty, M., & Khotilovich, V. (2019). Package 'xgboost'. R version, 90.

Chen, Z., Tao, R., Wu, X., Wei, Z., & Luo, X. (2019). Active learning for spam email classification. *ACM International Conference Proceeding Series*, 457–461. https://doi.org/10.1145/3377713.3377789

Chong, H. Y., Teoh, S. L., Wu, D. B. C., Kotirum, S., Chiou, C. F., & Chaiyakunapruk, N. (2016). Global economic burden of schizophrenia: A systematic review. In *Neuropsychiatric Disease and Treatment* (Vol. 12, pp. 357–373). Dove Medical Press Ltd. https://doi.org/10.2147/NDT.S96649

Coltheart, M., Langdon, R., & McKay, R. (2011). Delusional belief. *Annual Review of Psychology*, *62*, 271–298. https://doi.org/10.1146/annurev.psych.121208.131622

Costafreda, S. G., Brébion, G., Allen, P., McGuire, P. K., and Fu, C. H. Y. (2008). Affective modulation of external misattribution bias in source monitoring in schizophrenia. *Psychol. Med*. 38, 821–824. doi: 10.1017/S0033291708 003243

Correll, C. U., Galling, B., Pawar, A., Krivko, A., Bonetto, C., Ruggeri, M., Craig, T. J., Nordentoft, M., Srihari, V. H., Guloksuz, S., Hui, C. L. M., Chen, E. Y. H., Valencia, M., Juarez, F., Robinson, D. G., Schooler, N. R., Brunette, M. F., Mueser, K. T., Rosenheck, R. A., … Kane, J. M. (2018). Comparison of early intervention services vs treatment as usual for early-phase psychosis: A systematic review, meta-analysis, and meta-regression. *JAMA Psychiatry*. 75(6), 555–565). American Medical Association. https://doi.org/10.1001/jamapsychiatry.2018.0623

Coscia, D. M., Narr, K. L., Robinson, D. G., Hamilton, L. S., Sevy, S., Burdick, K. E., Gunduz-Bruce, H., McCormack, J., Bilder, R. M., & Szeszko, P. R. (2009). Volumetric and shape analysis of the thalamus in first-episode schizophrenia. *Human Brain Mapping*, *30*(4), 1236–1245. https://doi.org/10.1002/hbm.20595

Cui, L.-B., Cai, M., Wang, X.-R., Zhu, Y.-Q., Wang, L.-X., Xi, Y.-B., Wang, H.-N., Zhu, X., & Yin, H. (2019). Prediction of early response to overall treatment for schizophrenia: A functional magnetic resonance imaging study. *Brain and Behavior*, *9*(2). https://doi.org/10.1002/brb3.1211

Crow, T. J. (1980). Molecular pathology of schizophrenia: more than one disease process?. *British medical journal*, *280*(6207), 66.

Das, S., Dey, A., Pal, A., & Roy, N. (2015). Applications of artificial intelligence in machine learning: Review and Prospect. *International Journal of Computer Applications*, *115*(9), 31–41. https://doi.org/10.5120/20182-2402

Davatzikos C, Shen D, Gur RC, et al. Whole-Brain Morphometric Study of Schizophrenia Revealing a Spatially Complex Set of Focal Abnormalities. *Arch Gen Psychiatry.* 2005;62(11):1218–1227. doi:10.1001/archpsyc.62.11.1218

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A., & Leisch, F. (2021). e1071: Misc functions of the department of statistics, probability Theory Group (Formerly: E1071), TU Wien (R package version 1.7-9) [Computer software]. The Comprehensive R Archive Network. Available from https://CRAN.R-project.org/package=e107

Dean, D. J., Woodward, N., Walther, S., McHugo, M., Armstrong, K., & Heckers, S. (2020). Cognitive motor impairments and brain structure in schizophrenia spectrum disorder patients with a history of catatonia. *Schizophrenia Research*, *222*, 335–341. https://doi.org/10.1016/j.schres.2020.05.012

Deshmukh, S. S., & Knierim, J. J. (2012). Hippocampus. *Wiley Interdisciplinary Reviews: Cognitive Science,* 3(2), 231–251. https://doi.org/10.1002/wcs.1164

Díaz-Caneja, C. M., Alloza, C., Gordaliza, P. M., Fernández-Pena, A., De Hoyos, L., Santonja, J., ... & Janssen, J. (2021). Sex differences in lifespan trajectories and variability of human sulcal and gyral morphology. *Cerebral Cortex*, *31*(11), 5107-5120.

Dietterich, T. (1995). Overfitting and undercomputing in machine learning. *ACM computing surveys (CSUR)*, *27*(3), 326-327.

Dietterich, T. G. (2000). Ensemble Methods in Machine Learning. *Multiple Classifier Systems*, 1–15. https://doi.org/10.1007/3-540-45014-9_1

Dinga, R., Penninx, B. W. J. H., Veltman, D. J., Schmaal, L., & Marquand, A. F. (2019). Beyond accuracy: Measures for assessing machine learning models, pitfalls and guidelines. *BioRxiv*. https://doi.org/10.1101/743138

Dowle, M., & Srinivasan, A. (2021). data.table: Extension of `data.frame` (R package version 1.14.0) [Computer software]. The Comprehensive R Archive Network. Available from https://CRAN.R-project.org/package=data.table

Duan, K.-B., & Keerthi, S. S. (2005). Which Is the Best Multiclass SVM Method? An Empirical Study. *International workshop on multiple classifier system.* https://doi.org/10.1007/11494683_28

Elis, O., Caponigro, J. M., & Kring, A. M. (2013). Psychosocial treatments for negative symptoms in schizophrenia: Current practices and future directions. *Clinical Psychology Review*, 33(8), 914–928. https://doi.org/10.1016/j.cpr.2013.07.001

Emsley, R., Rabinowitz, J., Torreman, M., & RIS-INT-35 Early Psychosis Global Working Group (2003). The factor structure for the Positive and Negative Syndrome Scale (PANSS) in recent-onset psychosis. *Schizophrenia research*, *61*(1), 47–57. https://doi.org/10.1016/s0920-9964(02)00302-x

Fadel, S., Ghoniemy, S., Abdallah, M., Sorra, H. A., Ashour, A., & Ansary, A. (2016). Investigating the Effect of Different Kernel Functions on the Performance of SVM for Recognizing Arabic Characters. *IJACSA) International Journal of Advanced Computer Science and Applications*. 7(1). www.ijacsa.thesai.org

Fan, J., Upadhye, S., & Worster, A. (2006). Understanding receiver operating characteristic (ROC) curves. *Canadian Journal of Emergency Medicine*, 8(1), 19–20. https://doi.org/10.1017/S1481803500013336

Fan, Y., Shen, D., Gur, R. C., Gur, R. E., & Davatzikos, C. (2007). COMPARE: Classification of morphological patterns using adaptive regional elements. *IEEE Transactions on Medical Imaging,* 26(1), 93-105. doi:10.1109/TMI.2006.886812

Fervaha, G., Foussias, G., Agid, O., & Remington, G. (2014). Motivational and neurocognitive deficits are central to the prediction of longitudinal functional outcome in schizophrenia. *Acta Psychiatrica Scandinavica*, *130*(4), 290–299. https://doi.org/10.1111/acps.12289

Feyaerts, J., Henriksen, M. G., Vanheule, S., Myin-Germeys, I., & Sass, L. A. (2021). Delusions beyond beliefs: a critical overview of diagnostic, aetiological, and therapeutic schizophrenia research from a clinical-phenomenological perspective. In *The Lancet Psychiatry*. 8(3), 237–249. Elsevier Ltd. https://doi.org/10.1016/S2215-0366(20)30460-0

First, M., Botteron, K., Carter, C., Castellanos, F.X., Dickstein, D.P., Drevets, W., Kim, K.L., Pescosolido, M.F., Rausch, S., Seymour, K.E., Sheline, Y., Zubieta, J.K., (2012). Consensus Report of the APA Work Group on Neuroimaging Markers of Psychiatric Disorders. American Psychiatric Association.

Fischl B. (2012). FreeSurfer. *NeuroImage*, *62*(2), 774–781. https://doi.org/10.1016/j.neuroimage.2012.01.021

Fletcher, Paul & Frith, Chris. (2009). Perceiving is believing: A Bayesian approach to explaining the positive symptoms of schizophrenia. *Nature reviews Neuroscience*. 10. 48-58. 10.1038/nrn2536.

Fornito, A., Yücel, M., Dean, B., Wood, S. J., & Pantelis, C. (2009). Anatomical abnormalities of the anterior cingulate cortex in schizophrenia: Bridging the gap between neuroimaging and neuropathology. In *Schizophrenia Bulletin*, 35(5),973–993. https://doi.org/10.1093/schbul/sbn025

Freund, Y., & Schapire, R. E. (1996). Experiments with a new boosting algorithm. *icml*. 96, 148-156.

Friston, K. (2010). The free-energy principle: a unified brain theory?. *Nat Rev Neurosci* 11**,** 127–138. https://doi.org/10.1038/nrn2787

Frith, C. D. (1987). The positive and negative symptoms of schizophrenia reflect impairments in the perception and initiation of action. *Psychological medicine*, 17(3), 631-648.

Fujiwara, H., Hirao, K., Namiki, C., Yamada, M., Shimizu, M., Fukuyama, H., Hayashi, T., & Murai, T. (2007). Anterior cingulate pathology and social cognition in schizophrenia: A study of gray matter, white matter and sulcal morphometry. *NeuroImage*, *36*(4), 1236–1245. https://doi.org/10.1016/j.neuroimage.2007.03.068

Fujiwara, H., Shimizu, M., Hirao, K., Miyata, J., Namiki, C., Sawamoto, N., Fukuyama, H., Hayashi, T., & Murai, T. (2008). Female specific anterior cingulate abnormality and its association with empathic disability in schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, 32(7), 1728–1734. https://doi.org/10.1016/j.pnpbp.2008.07.013

Ganganwar, V. (2012). An overview of classification algorithms for imbalanced datasets. In *International Journal of Emerging Technology and Advanced Engineering*. 2(4). www.ijetae.com

Ghahramani Z. (2004) Unsupervised Learning. In: Bousquet O., von Luxburg U., Rätsch G. (eds) Advanced Lectures on Machine Learning. ML 2003. Lecture Notes in Computer Science, vol 3176. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-28650-9_5

Ghosh, S., Dasgupta, A., & Swetapadma, A. (2019). A study on support vector machine based linear and non-linear pattern classification. *2019 International Conference on Intelligent Sustainable Systems (ICISS).* https://doi.org/10.1109/iss1.2019.8908018

Giedd, J. N., Vaituzis, A. C., Hamburger, S. D., Lange, N., Rajapakse, J. C., Kaysen, D., Vauss, Y. C., & Rapoport, J. L. (1996). Quantitative MRI of the Temporal Lobe, Amygdala, and Hippocampus in Normal Human Development: Ages 4-18 Years. In *The Journal of Cognitive Biology* (Vol. 366).

Giraldo-Chica, M., Rogers, B. P., Damon, S. M., Landman, B. A., & Woodward, N. D. (2018). Prefrontal-Thalamic Anatomical Connectivity and Executive Cognitive Function in

Schizophrenia. *Biological Psychiatry*, *83*(6), 509–517. https://doi.org/10.1016/j.biopsych.2017.09.022

Glahn, D. C., Laird, A. R., Ellison-Wright, I., Thelen, S. M., Robinson, J. L., Lancaster, J. L., ... & Fox, P. T. (2008). Meta-analysis of gray matter anomalies in schizophrenia: application of anatomic likelihood estimation and network analysis. *Biological psychiatry*, *64*(9), 774-781. https://doi.org/10.1016/j.biopsych.2008.03.031

Gloutney, M. L., & Clark, R. G. (1991). The significance of body mass to female dabbling ducks during late incubation. https://doi.org/10.2307/3247715

Goff, D. C. (2021). The Pharmacologic Treatment of Schizophrenia (2021). *JAMA - Journal of the American Medical Association*, 325(2), 175–176. https://doi.org/10.1001/jama.2020.19048

Gong, J., Cui, L.-B., Xi, Y.-B., Zhao, Y.-S., Yang, X.-J., Xu, Z.-liang, Sun, J.-B., Liu, P., Jia, J., Li, P., Yin, H., & Qin, W. (2020). Predicting response to electroconvulsive therapy combined with antipsychotics in schizophrenia using multi-parametric Magnetic Resonance Imaging. *Schizophrenia Research*, *216*, 262–271. https://doi.org/10.1016/j.schres.2019.11.046

Grama, L., Tuns, L., & Rusu, C. (2017). On the optimization of SVM kernel parameters for improving audio classification accuracy. *2017 14th International Conference on Engineering of Modern Electric Systems, EMES 2017*, 224–227. https://doi.org/10.1109/EMES.2017.7980420

Greenstein, D., Malley, J. D., Weisinger, B., Clasen, L., & Gogtay, N. (2012). Using Multivariate Machine Learning Methods and Structural MRI to Classify Childhood Onset Schizophrenia and Healthy Controls. *Frontiers in Psychiatry*, *3*. https://doi.org/10.3389/fpsyt.2012.00053

Haijma, S. v., van Haren, N., Cahn, W., Koolschijn, P. C. M. P., Hulshoff Pol, H. E., & Kahn, R. S. (2013). Brain volumes in schizophrenia: A meta-analysis in over 18 000 subjects. *Schizophrenia Bulletin*, *39*(5), 1129–1138. https://doi.org/10.1093/schbul/sbs118

Hamelin, L., Bertoux, M., Bottlaender, M., Corne, H., Lagarde, J., Hahn, V., ... & Sarazin, M. (2015). Sulcal morphology as a new imaging marker for the diagnosis of early onset Alzheimer's disease. *Neurobiology of aging*, *36*(11), 2932-2939.

Hasan, A., Wolff-Menzler, C., Pfeiffer, S., Falkai, P., Weidinger, E., Jobst, A., Hoell, I., Malchow, B., Yeganeh-Doost, P., Strube, W., Quast, S., Müller, N., & Wobrock, T. (2015). Transcutaneous noninvasive vagus nerve stimulation (tVNS) in the treatment of schizophrenia: a bicentric randomized controlled pilot study. *European Archives of Psychiatry and Clinical Neuroscience*, *265*(7), 589–600. https://doi.org/10.1007/s00406-015-0618-9

Henssler, J., Brandt, L., Müller, M., Liu, S., Montag, C., Sterzer, P., & Heinz, A. (2020). Migration and schizophrenia: meta-analysis and explanatory framework. *European Archives of Psychiatry and Clinical Neuroscience*, *270*(3), 325–335. https://doi.org/10.1007/s00406-019-01028-7

Herbener, E. S. (2008). Emotional memory in schizophrenia. *Schizophrenia Bulletin, 34*(5), 875–887. https://doi.org/10.1093/schbul/sbn081

Hjern, A., Wicks, S., & Dalman, C. (2004). Social adversity contributes to high morbidity in psychoses in immigrants - A national cohort study in two generations of Swedish residents. *Psychological Medicine*, *34*(6), 1025–1033. https://doi.org/10.1017/S003329170300148X

Hirao, K., Miyata, J., Fujiwara, H., Yamada, M., Namiki, C., Shimizu, M., Sawamoto, N., Fukuyama, H., Hayashi, T., & Murai, T. (2008). Theory of mind and frontal lobe pathology in schizophrenia: A voxel-based morphometry study. *Schizophrenia Research*, *105*(1–3), 165–174. https://doi.org/10.1016/j.schres.2008.07.021

Ho, B. C., Andreasen, N. C., Nopoulos, P., Arndt, S., Magnotta, V., & Flaum, M. (2003). Progressive structural brain abnormalities and their relationship to clinical outcome: a longitudinal magnetic resonance imaging study early in schizophrenia. *Archives of general Psychiatry*, *60*(6), 585-594. doi:10.1001/archpsyc.60.6.585

Honea, R., Crow, T. J., Passingham, D., & Mackay, C. E. (2005). Regional deficits in brain volume in schizophrenia: a meta-analysis of voxel-based morphometry studies. *American Journal of Psychiatry*, *162*(12), 2233-2245. https://doi.org/10.1176/appi.ajp.162.12.2233

Hong, C. S., & Oh, T. G. (2021). TPR-TNR plot for confusion matrix. Communications for Statistical Applications and Methods, 28(2), 161–169. https://doi.org/10.29220/CSAM.2021.28.2.161

Hopkins, S. C., Ogirala, A., Loebel, A., & Koblan, K. S. (2018). Transformed PANSS Factors Intended to Reduce Pseudospecificity Among Symptom Domains and Enhance Understanding of Symptom Change in Antipsychotic-Treated Patients With Schizophrenia. *Schizophrenia Bulletin*, *44*(3), 593–602. https://doi.org/10.1093/schbul/sbx101

Howes, O., McCutcheon, R., & Stone, J. (2015). Glutamate and dopamine in schizophrenia: an update for the 21st century. *Journal of psychopharmacology*, *29*(2), 97-115. https://doi.org/10.1177%2F0269881114563634

Howes, O. D., McCutcheon, R., Agid, O., De Bartolomeis, A., Van Beveren, N. J., Birnbaum, M. L., ... & Correll, C. U. (2017). Treatment-resistant schizophrenia: treatment response and resistance in psychosis (TRRIP) working group consensus guidelines on diagnosis and terminology. *American Journal of Psychiatry*, *174*(3), 216-229. https://doi.org/10.1176/appi.ajp.2016.16050503

Hulshoff Pol, H. E.., Schnack, H. G., Mandl, R. C., van Haren, N. E., Koning, H., Collins, D. L., ... & Kahn, R. S. (2001). Focal gray matter density changes in schizophrenia. *Archives of General Psychiatry*, *58*(12), 1118-1125. doi:10.1001/archpsyc.58.12.1118

Hulshoff Pol, H. E., & Kahn, R. S. (2008). What happens after the first episode? A review of progressive brain changes in chronically ill patients with schizophrenia. *Schizophrenia Bulletin*, *34*(2), 354–366. https://doi.org/10.1093/schbul/sbm168

Hulshoff Pol, H. E., Schnack, H. G., Bertens, M. G., van Haren, N. E., van der Tweel, I., Staal, W. G., ... & Kahn, R. S. (2002). Volume changes in gray matter in patients with

schizophrenia. American Journal of Psychiatry, 159(2), 244-250. https://doi.org/10.1176/appi.ajp.159.2.244

Idanpaan-Heikkila J, Alhava E, Olkinuora M, Palva IP. (1977).Agranulocytosis during treatment with chlozapine.*Eur J Clin Pharmacol1977,*11(3) ,193–8. https://doi.org/10.1007/bf00606409

Iwata, Y., Nakajima, S., Suzuki, T., Keefe, R. S. E., Plitman, E., Chung, J. K., Caravaggio, F., Mimura, M., Graff-Guerrero, A., & Uchida, H. (2015). Effects of glutamate positive modulators on cognitive deficits in schizophrenia: A systematic review and meta-Analysis of double-blind randomized controlled trials. *Molecular Psychiatry*, 20(10), 1151–1160. Nature Publishing Group. https://doi.org/10.1038/mp.2015.68

Jääskeläinen, E., Juola, P., Hirvonen, N., McGrath, J. J., Saha, S., Isohanni, M., Veijola, J., & Miettunen, J. (2013). A systematic review and meta-analysis of recovery in Schizophrenia. *Schizophrenia Bulletin*, *39*(6), 1296–1306. https://doi.org/10.1093/schbul/sbs130

Jain, M., Narayan, S., Balaji, P., BharathK., P., Bhowmick, A., Karthik, R., & Muthu, R.K. (2020). Speech Emotion Recognition using Support Vector Machine. *ArXiv, abs/2002.07590*.

Jamadar, S., O'Neil, K. M., Pearlson, G. D., Ansari, M., Gill, A., Jagannathan, K., & Assaf, M. (2013). Impairment in semantic retrieval is associated with symptoms in schizophrenia but not bipolar disorder. *Biological Psychiatry*, *73*(6), 555–564. https://doi.org/10.1016/j.biopsych.2012.07.027

Janssen, J., Alemán-Gómez, Y., Schnack, H., Balaban, E., Pina-Camacho, L., Alfaro-Almagro, F., ... & Desco, M. (2014). Cortical morphology of adolescents with bipolar disorder and with schizophrenia. *Schizophrenia research*, *158*(1-3), 91-99. https://doi.org/10.1016/j.schres.2014.06.040

Janssen, J., Díaz-Caneja, C. M., Alloza, C., Schippers, A., De Hoyos, L., Santonja, J., ... & Schnack, H. G. (2021). Dissimilarity in sulcal width patterns in the cortex can be used to identify patients with schizophrenia with extreme deficits in cognitive performance. *Schizophrenia bulletin*, *47*(2), 552-561. https://doi.org/10.1093/schbul/sbaa131

Jilka, S., Odoi, C. M., van Bilsen, J., Morris, D., Erturk, S., Cummins, N., Cella, M., & Wykes, T. (2022). Identifying schizophrenia stigma on Twitter: a proof of principle model using service user supervised machine learning. *Npj Schizophrenia*, *8*(1), 1. https://doi.org/10.1038/s41537-021-00197-6

Jones, A. L., Mowry, B. J., Pender, M. P., & Greer, J. M. (2005). Immune dysregulation and self-reactivity in schizophrenia: Do some cases of schizophrenia have an autoimmune basis? *Immunology and Cell Biology*,83(1), 9–17. https://doi.org/10.1111/j.1440-1711.2005.01305.x

Johns, L. C., Gregg, L., Allen, P., and McGuire, P. K. (2006). Impaired verbal self monitoring in psychosis: effects of state, trait and diagnosis. Psychol. Med. 36, 465–474. doi: 10.1017/S0033291705006628

Johns, L. C., Rossell, S., Frith, C., Ahmad, F., Hemsley, D., Kuipers, E., et al. (2001). Verbal self-monitoring and auditory verbal hallucinations in patients with schizophrenia. Psychol. Med. 31, 705–715. doi: 10.1017/S0033291701003774

Kalmady, S. V., Shivakumar, V., Arasappa, R., Subramaniam, A., Gautham, S., Venkatasubramanian, G., & Gangadhar, B. N. (2017). Clinical correlates of hippocampus volume and shape in antipsychotic-naïve schizophrenia. *Psychiatry Research - Neuroimaging*, *263*, 93–102. https://doi.org/10.1016/j.pscychresns.2017.03.014

Kane, J. M., Agid, O., Baldwin, M. L., Howes, O., Lindenmayer, J. P., Marder, S., Olfson, M., Potkin, S. G., & Correll, C. U. (2019). Clinical guidance on the identification and management of treatment-resistant schizophrenia. *Journal of Clinical Psychiatry*, *80*(2). https://doi.org/10.4088/JCP.18com12123

Kay, S. R., Fiszbein, A., & Opler, L. A. (1987). The positive and negative syndrome scale (PANSS) for schizophrenia. *Schizophrenia Bulletin*, *13*(2), 261–276. https://doi.org/10.1093/schbul/13.2.261

Kegeles, L. S., Abi-Dargham, A., Frankle, W. G., Gil, R., Cooper, T. B., Slifstein, M., Hwang, D.-R., Huang, Y., Haber, S. N., & Laruelle, M. (2010). Increased synaptic dopamine function in associative regions of the striatum in Schizophrenia. *Archives of General Psychiatry*, *67*(3), 231. https://doi.org/10.1001/archgenpsychiatry.2010.10

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*. 17(1). BioMed Central Ltd. https://doi.org/10.1186/s12916-019-1426-2

Kempton, M. J., Stahl, D., Williams, S. C. R., & DeLisi, L. E. (2010). Progressive lateral ventricular enlargement in schizophrenia: A meta-analysis of longitudinal MRI studies. *Schizophrenia Research*, *120*(1–3), 54–62. https://doi.org/10.1016/j.schres.2010.03.036

Kerns, J. G., & Berenbaum, H. (2002). Cognitive impairments associated with formal thought disorder in people with schizophrenia. *Journal of Abnormal Psychology*, *111*(2), 211–224. https://doi.org/10.1037/0021-843X.111.2.211

Khosla, N. (2004). Dimensionality reduction using factor analysis. Griffith University, Australia.

Killgore, W. D. S., Rosso, I. M., Gruber, S. A., & Yurgelun-Todd, D. A. (2009). *Amygdala Volume and Verbal Memory Performance in Schizophrenia and Bipolar Disorder*. http://journals.lww.com/cogbehavneurol

Kim, J. J., Crespo-Facorro, B., Andreasen, N. C., O'Leary, D. S., Magnotta, V., & Nopoulos, P. (2003). Morphology of the lateral superior temporal gyrus in neuroleptic naïve patients with schizophrenia: Relationship to symptoms. *Schizophrenia Research*, *60*(2–3), 173–181. https://doi.org/10.1016/S0920-9964(02)00299-2

Kim, J. H., Kim, S. Y., Lee, J., Oh, K. J., Kim, Y. B., & Cho, Z. H. (2012). Evaluation of the factor structure of symptoms in patients with schizophrenia. *Psychiatry research*, *197*(3), 285-289.

Kirkpatrick, B. K. (2014). Developing concepts in negative symptoms: Primary vs secondary and apathy vs expression. *Journal of Clinical Psychiatry*, *75*(SUPPL. 1), 3–7. https://doi.org/10.4088/JCP.13049su1c.01

Klein, M. O., Battagello, D. S., Cardoso, A. R., Hauser, D. N., Bittencourt, J. C., & Correa, R. G. (2019). Dopamine: Functions, Signaling, and Association with Neurological Diseases. *Cellular and Molecular Neurobiology*. 39(1), pp. 31–59). Springer New York LLC. https://doi.org/10.1007/s10571-018-0632-3

Klein, A., Ghosh, S. S., Bao, F. S., Giard, J., Häme, Y., Stavsky, E., ... & Keshavan, A. (2017). Mindboggling morphometry of human brains. *PLoS computational biology*, *13*(2), e1005350. https://doi.org/10.1371/journal.pcbi.1005350

Kochunov, P., Robin, D. A., Royall, D. R., Coyle, T., Lancaster, J., Kochunov, V., ... & Fox, P. T. (2009). Can structural MRI indices of cerebral integrity track cognitive trends in executive control function during normal maturation and adulthood?. *Human brain mapping*, *30*(8), 2581-2594. https://doi.org/10.1002/hbm.20689

Kodinariya, T. M., & Makwana, P. R. (2013). Review on determining number of Cluster in K-Means Clustering. *International Journal*, 1(6), 90-95.

Koutsouleris, N., Gaser, C., Jäger, M., Bottlender, R., Frodl, T., Holzinger, S., Schmitt, G. J. E., Zetzsche, T., Burgermeister, B., Scheuerecker, J., Born, C., Reiser, M., Möller, H. J., & Meisenzahl, E. M. (2008). Structural correlates of psychopathological symptom dimensions in schizophrenia: A voxel-based morphometric study. *NeuroImage*, *39*(4), 1600–1612. https://doi.org/10.1016/j.neuroimage.2007.10.029

Korver, N., Quee, P. J., Boos, H. B., Simons, C. J., de Haan, L., & Group Investigators. (2012). Genetic Risk and Outcome of Psychosis (GROUP), a multi site longitudinal cohort study focused on gene–environment interaction: objectives, sample characteristics, recruitment

and assessment methods. *International journal of methods in psychiatric research*, *21*(3), 205-221. https://doi.org/10.1002/mpr.1352

Kring, A. M., & Moran, E. K. (2008). Emotional response deficits in schizophrenia: Insights from affective science. In *Schizophrenia Bulletin*.34(5), 819–834. https://doi.org/10.1093/schbul/sbn071

Kubota, M., van Haren, N. E., Haijma, S. V., Schnack, H. G., Cahn, W., Pol, H. E. H., & Kahn, R. S. (2015). Association of IQ changes and progressive brain changes in patients with schizophrenia. *JAMA psychiatry*, *72*(8), 803-812. doi:10.1001/jamapsychiatry.2015.0712

Kuhn, M. (2021). caret: Classification and regression training (R package version 6.0-90) [Computer software]. The Comprehensive R Archive Network. Available from https://CRAN.R-project.org/package=caret

Kulynych, J. J., Luevano, L. F., Jones, D. W., & Weinberger, D. R. (1997). Cortical abnormality in schizophrenia: an in vivo application of the gyrification index. *Biological psychiatry*, *41*(10), 995-999.

Kung, C., & Yu, R. (2020). Interpretable models do not compromise accuracy or fairness in predicting college success. In Proceedings of the seventh acm conference on learning scale (413-416). https://doi.org/10.1145/3386527.3406755

Lambert, T. J. R., Velakoulis, D., & Pantelis, C. (2003). Medical comorbidity in schizophrenia. *MJA*, 178. http://www.mja.com.au/public/issues/172_12_190600/

Langdon, R., & Coltheart, M. (2000). The cognitive neuropsychology of delusions. *Mind & Language,* 15, 183–216. https://doi.org/10.1111/ 1468-0017.00129

Laursen, T. M., Munk-Olsen, T., & Vestergaard, M. (2012). Life expectancy and cardiovascular mortality in persons with schizophrenia. In *Current Opinion in Psychiatry*,25(2), 83–88. https://doi.org/10.1097/YCO.0b013e32835035ca

LaValley, M. P. (2008). Logistic regression. *Circulation*, 117(18), 2395-2399. https://doi.org/10.1161/CIRCULATIONAHA.106.682658

LeDoux, J. (2007). The amygdala. *Current Biology*, 17(20). https://doi.org/10.1016/j.cub.2007.08.005

Leow, A. D., Zhan, L., Zhu, S., Hageman, N., Chiang, M. C., Barysheva, M., Toga, A. W., McMahon, K. L., de Zubicaray, G. I., Wright, M. J., & Thompson, P. M. (2009). White matter integrity measured by fractional anisotropy correlates poorly with actual individual fiber anisotropy. *Proceedings - 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2009*, 622–625. https://doi.org/10.1109/ISBI.2009.5193124

Lesh, T. A., Tanase, C., Geib, B. R., Niendam, T. A., Yoon, J. H., Minzenberg, M. J., Ragland, J. D., Solomon, M., & Carter, C. S. (2015). A multimodal analysis of antipsychotic effects on brain structure and function in first-episode schizophrenia. *JAMA Psychiatry*, *72*(3), 226–234. https://doi.org/10.1001/jamapsychiatry.2014.2178

Leucht, S., Burkard, T., Henderson, J., Maj, M., & Sartorius, N. (2007). Physical illness and schizophrenia: a review of the literature. *Acta Psychiatrica Scandinavica*, 116(5), 317-333. https://doi.org/10.1111/j.1600-0447.2007.01095.x

Levine, S. Z., & Rabinowitz, J. (2007). Revisiting the 5 dimensions of the Positive and Negative Syndrome Scale. *Journal of clinical psychopharmacology*, 27(5), 431-436. doi: 10.1097/jcp/.0b013e31814cfabd

Lewis, D. A., & Moghaddam, B. (2006). Cognitive dysfunction in schizophrenia. *Archives of Neurology*, *63*(10), 1372. https://doi.org/10.1001/archneur.63.10.1372

Liaw A, Wiener M (2002). "Classification and Regression by randomForest." R News, 2(3), 18-22. https://CRAN.R-project.org/doc/Rnews/.

Lieberman, J. A., Stroup, T. S., Perkins, D. O., & Dixon, L. B. (2020). *The American Psychiatric Association Publishing Textbook of Schizophrenia*. American Psychiatric Association Publishing.

Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern recognition*, 36(2), 451-461.

Lipner, E., Murphy, S. K., & Ellman, L. M. (2019). Prenatal Maternal Stress and the Cascade of Risk to Schizophrenia Spectrum Disorders in Offspring. In *Current Psychiatry Reports*, 21(10). Current Medicine Group LLC 1. https://doi.org/10.1007/s11920-019-1085-1

Liu, T., Wen, W., Zhu, W., Kochan, N. A., Trollor, J. N., Reppermund, S., ... & Sachdev, P. S. (2011). The relationship between cortical sulcal variability and cognitive performance in

the elderly. *Neuroimage*, 56(3), 865-873. https://doi.org/10.1016/j.neuroimage.2011.03.015

Luo, Y., Zhang, L., He, P., Pang, L., Guo, C., & Zheng, X. (2019). Individual-level and area-level socioeconomic status (SES) and schizophrenia: Cross-sectional analyses using the evidence from 1.9 million Chinese adults. *BMJ Open*, *9*(9). https://doi.org/10.1136/bmjopen-2018-026532

Luther, L., Fukui, S., Firmin, R. L., McGuire, A. B., White, D. A., Minor, K. S., & Salyers, M. P. (2015). Expectancies of success as a predictor of negative symptoms reduction over 18 months in individuals with schizophrenia. Psychiatry Research, 229(1–2), 505–510. https://doi.org/10.1016/j.psychres.2015.06.022

Ljubobratović, D., & Matetić, M. (2019). Using LMS Activity Logs to Predict Student Failure with Random Forest Algorithm. *The Future of Information Sciences*, 113.

Makin, T. R., & Flor, H. (2020). Brain (re)organisation following amputation: Implications for phantom limb pain. NeuroImage, 218. https://doi.org/10.1016/j.neuroimage.2020.116943

Marder, S. R., & Galderisi, S. (2017). The current conceptualization of negative symptoms in schizophrenia. *World Psychiatry*, *16*(1), 14–24. https://doi.org/10.1002/wps.20385

Mangin, J. F., Riviere, D., Cachia, A., Duchesnay, E., Cointepas, Y., Papadopoulos-Orfanos, D., ... & Régis, J. (2004). A framework to study the cortical folding patterns. *Neuroimage*, *23*, 129-138. https://doi.org/10.1016/j.neuroimage.2004.07.019

Maulud, D., & Abdulazeez, A. M. (2020). A Review on Linear Regression Comprehensive in Machine Learning. *Journal of Applied Science and Technology Trends*, *1*(4), 140–147. https://doi.org/10.38094/jastt1457

Matsuda, Y., & Ohi, K. (2018). Cortical gyrification in schizophrenia: current perspectives. *Neuropsychiatric Disease and Treatment*, *14*, 1861.

McCleery, A., Wynn, J. K., & Green, M. F. (2018). Hallucinations, neuroplasticity, and prediction errors in schizophrenia. *Scandinavian Journal of Psychology*, *59*(1), 41–48. https://doi.org/10.1111/sjop.12413

McGrath, J., Saha, S., Welham, J., El Saadi, O., MacCauley, C., & Chant, D. (2004). A systematic review of the incidence of schizophrenia: The distribution of rates and the influence of sex, urbanicity, migrant status and methodology. *BMC Medicine*, *2*(1). https://doi.org/10.1186/1741-7015-2-13

McKinney, B., Ding, Y., Lewis, D. A., & Sweet, R. A. (2017). DNA methylation as a putative mechanism for reduced dendritic spine density in the superior temporal gyrus of subjects with schizophrenia. *Translational Psychiatry*, *7*(2). https://doi.org/10.1038/tp.2016.297

Meduri, M., Bramanti, P., Ielitro, G., Favaloro, A., Milardi, D., Cutroneo, G., Muscatello, M. R. A., Bruno, A., Micò, U., Pandolfo, G., la Torre, D., Vaccarino, G., & Anastasi, G. (2010). Morphometrical and morphological analysis of lateral ventricles in schizophrenia patients versus healthy controls. *Psychiatry Research - Neuroimaging*, *183*(1), 52–58. https://doi.org/10.1016/j.pscychresns.2010.01.014

Menoret, M., Farrugia, N., Pasdeloup, B., & Gripon, V. (2018). Evaluating graph signal processing for neuroimaging through classification and dimensionality reduction. *2017 IEEE Global Conference on Signal and Information Processing, GlobalSIP 2017 - Proceedings*, *2018-January*, 618–622. https://doi.org/10.1109/GlobalSIP.2017.8309033

Mijovic, A., & MacCabe, J. H. (2020). Clozapine-induced agranulocytosis. *Annals of Hematology*, 99(11), 2477–2482. https://doi.org/10.1007/s00277-020-04215-y

Minzenberg, M. J., Firl, A. J., Yoon, J. H., Gomes, G. C., Reinking, C., & Carter, C. S. (2010). Gamma oscillatory power is impaired during cognitive control independent of medication status in first-episode schizophrenia. *Neuropsychopharmacology*, *35*(13), 2590–2599. https://doi.org/10.1038/npp.2010.150

Minzenberg, M. J., Lesh, T. A., Niendam, T. A., Yoon, J. H., Rhoades, R. N., & Carter, C. S. (2014). Frontal cortex control dysfunction related to long-term suicide risk in recent-onset schizophrenia. *Schizophrenia Research*, *157*(1–3), 19–25. https://doi.org/10.1016/j.schres.2014.05.039

Moghaddam, B. (2003). Bringing order to the glutamate chaos in schizophrenia. *Neuron*, *40*(5), 881–884. https://doi.org/10.1016/s0896-6273(03)00757-8

Moghaddam, B., & Javitt, D. (2012). From revolution to evolution: The glutamate hypothesis of schizophrenia and its implication for treatment. *Neuropsychopharmacology,* 37(1), 4–15. https://doi.org/10.1038/npp.2011.181

Moskowitz, A. (2008). Association and dissociation in the historical concept of schizophrenia. In A. Moskowitz, I. Schäfer, & M. J. Dorahy (Eds.), *Psychosis, trauma and dissociation:*

Emerging perspectives on severe psychopathology. 35–49. Wiley-Blackwell. https://doi.org/10.1002/9780470699652.ch3

Mouselimis, L. (2021). {ClusterR}: Gaussian mixture models, k-means, mini-batch-kmeans, k-medoids and affinity propagation clustering (R package version 1.2.5) [Computer software]. The Comprehensive R Archive Network. Available from https://CRAN.R-project.org/package=ClusterR

Mubarik, A., & Tohid, H. (2016). Frontal lobe alterations in schizophrenia: A review. In *Trends in Psychiatry and Psychotherapy* ,38(4),198–206. https://doi.org/10.1590/2237-6089-2015-0088

Mueser, K. T., & Jeste, D. V. (2008). *Clinical handbook of schizophrenia*. Guilford Press.

Mueser, K. T., Deavers, F., Penn, D. L., & Cassisi, J. E. (2013). Psychosocial treatments for schizophrenia. In *Annual Review of Clinical Psychology* ,9,465–497. https://doi.org/10.1146/annurev-clinpsy-050212-185620

Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences*, *116*(44), 22071–22080. https://doi.org/10.1073/pnas.1900654116

Najjar, S., & Pearlman, D. M. (2015). Neuroinflammation and white matter pathology in schizophrenia: Systematic review. In *Schizophrenia Research*,161(1), 102–112. https://doi.org/10.1016/j.schres.2014.04.041

Nasar S. A beautiful mind. New York: Touchstone (Simon & Schuster); 1998.

Nasr, T., & Kausar, R. (2009). Psychoeducation and the family burden in schizophrenia: A randomized controlled trial. Annals of General Psychiatry, 8(1). https://doi.org/10.1186/1744-859x-8-17

Nenadic, I., Sauer, H., & Gaser, C. (2010). Distinct pattern of brain structural deficits in subsyndromes of schizophrenia delineated by psychopathology. *NeuroImage*, *49*(2), 1153–1160. https://doi.org/10.1016/j.neuroimage.2009.10.014

Nieuwenhuis, M., van Haren, N. E. M., Hulshoff Pol, H. E., Cahn, W., Kahn, R. S., & Schnack, H. G. (2012). Classification of schizophrenia patients and healthy controls from structural MRI scans in two large independent samples. NeuroImage, 61(3), 606–612. https://doi.org/10.1016/j.neuroimage.2012.03.079

Nielsen, D. (2016). *Tree boosting with xgboost-why does xgboost win" every" machine learning competition?* (Master's thesis, NTNU).

Niu, L., Matsui, M., Zhou, S. Y., Hagino, H., Takahashi, T., Yoneyama, E., Kawasaki, Y., Suzuki, M., Seto, H., Ono, T., & Kurachi, M. (2004). Volume reduction of the amygdala in patients with schizophrenia: A magnetic resonance imaging study. *Psychiatry Research - Neuroimaging*, *132*(1), 41–51. https://doi.org/10.1016/j.pscychresns.2004.06.002

Niznikiewicz, M. (2008). Future directions for examining semantic memory in schizophrenia spectrum disorders. *Clinical EEG and Neuroscience*, *39*(2), 95–98. https://doi.org/10.1177/155005940803900215

Niznikiewicz, M., Mittal, M. S., Nestor, P. G., & McCarley, R. W. (2010). Abnormal inhibitory processes in semantic networks in schizophrenia. *International Journal of Psychophysiology*, 75(2), 133–140. https://doi.org/10.1016/j.ijpsycho.2009.10.006

Ohi, K., Matsuda, Y., Shimada, T., Yasuyama, T., Oshima, K., Sawai, K., Kihara, H., Nitta, Y., Okubo, H., Uehara, T., & Kawasaki, Y. (2016). Structural alterations of the superior temporal gyrus in schizophrenia: Detailed subregional differences. *European Psychiatry*, *35*, 25–31. https://doi.org/10.1016/j.eurpsy.2016.02.002

Park, H. A. (2013). An introduction to logistic regression: From basic concepts to interpretation with particular attention to nursing domain. *Journal of Korean Academy of Nursing*, *43*(2), 154–164. https://doi.org/10.4040/jkan.2013.43.2.154

Pesantez-Narvaez, J., Guillen, M., & Alcañiz, M. (2019). Predicting motor insurance claims using telematics data—XGboost versus logistic regression. *Risks*, *7*(2). https://doi.org/10.3390/risks7020070

Pettersson-Yeo, W., Benetti, S., Marquand, A., Dell'Acqua, F., Williams, S., Allen, P., . . . Mechelli, A. (2013). Using genetic, cognitive and multi-modal neuroimaging data to identify ultra-high-risk and first-episode psychosis at the individual level. *Psychological Medicine, 43*(12), 2547-2562. doi:10.1017/S003329171300024X

Pillny, M., Schlier, B., & Lincoln, T. M. (2020). "I just don't look forward to anything". How anticipatory pleasure and negative beliefs contribute to goal-directed activity in patients with negative symptoms of psychosis. *Schizophrenia Research*, *222*, 429–436. https://doi.org/10.1016/j.schres.2020.03.059

Phillips, J., Cripps, E., Lau, J. W., & Hodkiewicz, M. R. (2015). Classifying machinery condition using oil samples and binary logistic regression. *Mechanical Systems and Signal Processing*, *60*, 316-325. https://doi.org/10.1016/j.ymssp.2014.12.020

Probst, P., & Boulesteix, A.-L. (2018). To Tune or Not to Tune the Number of Trees in Random Forest. In *Journal of Machine Learning Research* (Vol. 18). http://jmlr.org/papers/v18/17-269.html.

Poulet F. (2005). VISUAL SVM. *Proceedings of the Seventh International Conference on Enterprise Information Systems*, 309-31.doi: 10.5220/0002521003090314

Pugliese, V., Bruni, A., Carbone, E. A., Calabrò, G., Cerminara, G., Sampogna, G., Luciano, M., Steardo, L., Fiorillo, A., Garcia, C. S., & de Fazio, P. (2019). Maternal stress, prenatal medical illnesses and obstetric complications: Risk factors for schizophrenia spectrum disorder, bipolar disorder and major depressive disorder. *Psychiatry Research*, *271*, 23–30. https://doi.org/10.1016/j.psychres.2018.11.023

Pfefferbaum, A., Lim, K. O., Zipursky, R. B., Mathalon, D. H., Rosenbloom, M. J., Lane, B., ... & Sullivan, E. V. (1992). Brain gray and white matter volume loss accelerates with aging in chronic alcoholics: a quantitative MRI study. Alcoholism: Clinical and Experimental Research, 16(6), 1078-1089.

Qiu, A., Zhong, J., Graham, S., Chia, M. Y., & Sim, K. (2009). Combined analyses of thalamic volume, shape and white matter integrity in first-episode schizophrenia. *NeuroImage*, *47*(4), 1163–1171. https://doi.org/10.1016/j.neuroimage.2009.04.027

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org/.

Rajarethinam, R. P., Dequardo, J. R., Nalepa, R., & Tandon, R. (2000). Superior temporal gyrus in schizophrenia: a volumetric magnetic resonance imaging study. In *Schizophrenia Research*. 41. www.elsevier.com/locate/schres

Ratti, M. T., Bo, P., Giardini, A., & Soragna, D. (2002). Chronic alcoholism and the frontal lobe: Which executive functions are imparied*? Acta Neurologica Scandinavica*, 105(4), 276–281. https://doi.org/10.1034/j.1600-0404.2002.0o315.x

Razzouk, D., Mari, J. D. J., Shirakawa, I., Wainer, J., & Sigulem, D. (2006). Decision support system for the diagnosis of schizophrenia disorders. *Brazilian journal of medical and biological research*, *39*(1), 119-128.

Rector, N. A., Beck, A. T., & Stolar, N. (2005). The negative symptoms of schizophrenia: A cognitive perspective. *The Canadian Journal of Psychiatry*, *50*(5), 247–257. https://doi.org/10.1177/070674370505000503

Ren, W., Su Lui, M., Deng, W., Li, F., Li, M., Huang, X., Wang, Y., Li, T., Sweeney, J. A., & Gong, Q. (2013). Anatomical and Functional Brain Abnormalities in Drug-Naive First-Episode Schizophrenia. *Am J Psychiatr,* 170(11).1308-1316-http://www.fil.ion.ucl.ac.uk/spm

Rich, A. M., Cho, Y. T., Tang, Y., Savic, A., Krystal, J. H., Wang, F., Xu, K., & Anticevic, A. (2016). Amygdala volume is reduced in early course schizophrenia. *Psychiatry Research - Neuroimaging*, *250*, 50–60. https://doi.org/10.1016/j.pscychresns.2016.02.006

Ripke, S., Neale, B. M., Corvin, A., Walters, J. T. R., Farh, K. H., Holmans, P. A., Lee, P., Bulik-Sullivan, B., Collier, D. A., Huang, H., Pers, T. H., Agartz, I., Agerbo, E., Albus, M., Alexander, M., Amin, F., Bacanu, S. A., Begemann, M., Belliveau, R. A., … O'Donovan, M. C. (2014). Biological insights from 108 schizophrenia-associated genetic loci. Nature, 511(7510), 421–427. https://doi.org/10.1038/nature13595

Riquelme, J. C., Ruiz, R., Rodríguez, D., & Moreno, J. (2008). Finding defective modules from highly unbalanced datasets. *Actas de los Talleres de las Jornadas de Ingeniería del Software y Bases de Datos*, 2(1), 67-74.

Rizopoulos, D. (2006). ltm: An r package for latent variable modelling and item response theory analyses () [Computer software]. The Comprehensive R Archive Network. Available from https://doi.org/10.18637/jss.v017.i05

Romine, C. B., & Reynolds, C. R. (2005). A model of the development of frontal lobe functioning: Findings from a meta-analysis. In *Applied Neuropsychology*, 12 (4),190–201. https://doi.org/10.1207/s15324826an1204_2

Roopa, H., & Asha, T. (2019). A Linear Model Based on Principal Component Analysis for Disease Prediction. *IEEE Access*, *7*, 105314–105318. https://doi.org/10.1109/ACCESS.2019.2931956

Chen, T., He, T., Benesty, M., Khotilovich, V., Tang, Y., Cho, H., Chen, K., Mitchell, R., Cano, I., Zhou, T., Li, M., Xie, J., Lin, M., Geng, Y., & Li, Y. (2021). xgboost: Extreme gradient boosting (R package version 1.5.0.2) [Computer software]. The Comprehensive R Archive Network. Available from https://CRAN.R-project.org/package=xgboost

Rosseel, Y., (2012). lavaan: An R Package for Structural Equation Modeling. Journal of Statistical Software, 48(2), 1-36. https://doi.org/10.18637/jss.v048.i02

Roth, R. M., Flashman, L. A., Saykin, A. J., Thomas McAllister, P. W., & Vidaver, R. (2004). Brief Report Apathy in Schizophrenia: Reduced Frontal Lobe Volume and Neuropsychological Deficits. *Am J Psychiatry*.161. http://ajp.psychiatryonline.org

Saha, S., Chant, D., Welham, J., & McGrath, J. (2005). A systematic review of the prevalence of schizophrenia. In *PLoS Medicine*, 2(5), 0413–0433. https://doi.org/10.1371/journal.pmed.0020141

Saifudin, A., Heryadi, Y., & Lukas. (2019). Ensemble Undersampling to Handle Unbalanced Class on Cross-Project Defect Prediction. IOP Conference Series: Materials Science and Engineering, 662(6). https://doi.org/10.1088/1757-899X/662/6/062012

Salgado-Pineda, P., Radua, J., Sarró, S., Guerrero-Pedraza, A., Salvador, R., Pomarol-Clotet, E., & McKenna, P. J. (2018). Sensitivity and specificity of hypoactivations and failure of de-activation in schizophrenia. Schizophrenia Research, 201, 224–230. https://doi.org/10.1016/j.schres.2018.06.013

Salomon JA, Vos T, Hogan DR, Gagnon M, Naghavi M, Mokdad A, Begum N…Jonas JB. Common values in assessing health outcomes from disease and injury: disability weights

measurement study for the Global Burden of Disease Study 2010. *Lancet*. 2012 Dec 15;380(9859):2129-43. doi: 10.1016/S0140-6736(12)61680-8.

Sampogna, G., Vecchio, V. del, Giallonardo, V., Luciano, M., Perris, F., Saviano, P., Zinno, F., & Fiorillo, A. (2020). Il processo di revisione dei sistemi diagnostici in psichiatria: differenze tra ICD-11 e DSM-5 The revision process of diagnostic systems in psychiatry: differences between ICD-11 and DSM-5. In *Riv Psichiatr*, 55(6).

Samuel, R., Thomas, E., & Jacob, K. S. (2018). Instrumental activities of daily living dysfunction among people with schizophrenia. *Indian Journal of Psychological Medicine*, *40*(2), 134–138. https://doi.org/10.4103/IJPSYM.IJPSYM_308_17

Saravanan R. &, Sujatha P. (2018). A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification, *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2018, pp. 945-949, doi: 10.1109/ICCONS.2018.8663155.

Sarmento, R. P., & Costa, V. (2019). Confirmatory factor analysis--a case study. *arXiv preprint arXiv:1905.05598*.

Scahill, R. I., Frost, C., Jenkins, R., Whitwell, J. L., Rossor, M. N., & Fox, N. C. (2003). A Longitudinal Study of Brain Volume Changes in Normal Aging Using Serial Registered Magnetic Resonance Imaging. In *Arch Neurol* (Vol. 60). https://jamanetwork.com/

Schnack, H. G. (2017). Improving individual predictions: Machine learning approaches for detecting and attacking heterogeneity in schizophrenia (and other psychiatric diseases). *Schizophrenia Research*, *214*, 34–42. https://doi.org/10.1016/j.schres.2017.10.023

Schnack, H. G., & Kahn, R. S. (2016). Detecting neuroimaging biomarkers for psychiatric disorders: Sample size matters. Frontiers in Psychiatry, 7. https://doi.org/10.3389/fpsyt.2016.00050

Schnell, K., Heekeren, K., Daumann, J., Schnell, T., Schnitker, R., Möller-Hartmann, W., & Gouzoulis-Mayfrank, E. (2008). Correlation of passivity symptoms and dysfunctional visuomotor action monitoring in psychosis. *Brain*, *131*(10), 2783–2797. https://doi.org/10.1093/brain/awn184

Seeman P. (2009) Glutamate and dopamine components in schizophrenia. *J Psychiatry Neurosci*. 34, 143–149.

Shepherd, A. M., Quidé, Y., Laurens, K. R., O'Reilly, N., Rowland, J. E., Mitchell, P. B., Vaughan, C. J., & Green, M. J. (2015). Shared intermediate phenotypes for schizophrenia and bipolar disorder: Neuroanatomical features of subtypes distinguished by executive dysfunction. *Journal of Psychiatry and Neuroscience*, *40*(1), 58–68. https://doi.org/10.1503/jpn.130283

Sheridan, R. P., Wang, W. M., Liaw, A., Ma, J., & Gifford, E. M. (2016). Extreme Gradient Boosting as a Method for Quantitative Structure-Activity Relationships. *Journal of Chemical Information and Modeling*, *56*(12), 2353–2360. https://doi.org/10.1021/acs.jcim.6b00591

Singh A., Thakur N. & Sharma A. (2016). A review of supervised machine learning algorithms, *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom),* pp. 1310-1315.

Siskind, D., McCartney, L., Goldschlager, R., & Kisely, S. (2016). Clozapine v . first- and second-generation antipsychotics in treatment-refractory schizophrenia: systematic review and meta-analysis . *British Journal of Psychiatry*, *209*(5), 385–392. https://doi.org/10.1192/bjp.bp.115.177261

Solmi, M., Murru, A., Pacchiarotti, I., Undurraga, J., Veronese, N., Fornaro, M., Stubbs, B., Monaco, F., Vieta, E., Vseeman, M., Correll, C. U., & Carvalho, A. F. (2017). Safety, tolerability, and risks associated with first-and second-generation antipsychotics: A state-of-the-art clinical review. *Therapeutics and Clinical Risk Management,* 13, 757–777. https://doi.org/10.2147/TCRM.S117321

Song, J., Han, D. H., Kim, S. M., Hong, J. S., Min, K. J., Cheong, J. H., & Kim, B. N. (2015). Differences in gray matter volume corresponding to delusion and hallucination in patients with schizophrenia compared with patients who have bipolar disorder. *Neuropsychiatric Disease and Treatment*, *11*, 1211–1219. https://doi.org/10.2147/NDT.S80438

Sorkin, A., Weinshall, D., Modai, I., & Peled, A. (2006). Improving the accuracy of the diagnosis of schizophrenia by means of virtual reality. *American Journal of Psychiatry*, *163*(3), 512–520. https://doi.org/10.1176/appi.ajp.163.3.512

Souli, S., & Lachiri, Z. (2018). Audio sounds classification using scattering features and support vectors machines for medical surveillance. *Applied Acoustics*, *130*, 270–282. https://doi.org/10.1016/j.apacoust.2017.08.002

Spoletini, I., Piras, F., Fagioli, S., Rubino, I. A., Martinotti, G., Siracusano, A., Caltagirone, C., & Spalletta, G. (2011). Suicidal attempts and increased right amygdala volume in

schizophrenia. *Schizophrenia Research*, *125*(1), 30–40. https://doi.org/10.1016/j.schres.2010.08.023

Srivastava, P. K., Han, D., Rico-Ramirez, M. A., Bray, M., & Islam, T. (2012). Selection of classification techniques for land use/land cover change investigation. *Advances in Space Research*, *50*(9), 1250–1265. https://doi.org/10.1016/j.asr.2012.06.032

Sterzer, P., Adams, R. A., Fletcher, P., Frith, C., Lawrie, S. M., Muckli, L., Petrovic, P., Uhlhaas, P., Voss, M., & Corlett, P. R. (2018). The Predictive Coding Account of Psychosis. In *Biological Psychiatry*. 84(9), 634–643. Elsevier USA. https://doi.org/10.1016/j.biopsych.2018.05.015

Stilo, S. A., & Murray, R. M. (2019). Non-Genetic Factors in Schizophrenia. In Current Psychiatry Reports 21(10). Current Medicine Group LLC 1. https://doi.org/10.1007/s11920-019-1091-3

Stirling, J., Hellewell, J., Blakey, A., & Deakin, W. (2006). Thought disorder in schizophrenia is associated with both executive dysfunction and circumscribed impairments in semantic function. *Psychological Medicine*, *36*(4), 475–484. https://doi.org/10.1017/S0033291705006884

Stoltzfus, J. C. (2011). Logistic regression: A brief primer. Academic Emergency Medicine, 18(10), 1099–1104. https://doi.org/10.1111/j.1553-2712.2011.01185.x

Strauss, G. P., Waltz, J. A., & Gold, J. M. (2014). A review of reward processing and motivational impairment in schizophrenia. In *Schizophrenia Bulletin* (Vol. 40, Issue SUPPL. 2). Oxford University Press. https://doi.org/10.1093/schbul/sbt197

Sun, J., Maller, J. J., Guo, L., & Fitzgerald, P. B. (2009). Superior temporal gyrus volume change in schizophrenia: A review on Region of Interest volumetric studies. In *Brain Research Reviews*, 61(1), 14–32. https://doi.org/10.1016/j.brainresrev.2009.03.004

Sundaram, R.S., Gowtham, L., & Nayak, B. S. (2012). The role of excitatory neurotransmitter glutamate in brain physiology and pathology *Asian Journal of Pharmaceutical and Clinical Research*,5,1-7.

Svancer, P., & Spaniel, F. (2021). Brain ventricular volume changes in schizophrenia. A narrative review. *Neuroscience Letters*, *759*. https://doi.org/10.1016/j.neulet.2021.136065

Talpalaru, A., Bhagwat, N., Devenyi, G. A., Lepage, M., & Chakravarty, M. M. (2019). Identifying schizophrenia subgroups using clustering and supervised learning. *Schizophrenia Research*, *214*, 51–59. https://doi.org/10.1016/j.schres.2019.05.044

Tan, E. J., Meyer, D., Neill, E., & Rossell, S. L. (2021). Investigating the diagnostic utility of speech patterns in schizophrenia and their symptom associations. *Schizophrenia Research*, *238*, 91–98. https://doi.org/10.1016/j.schres.2021.10.003

Tan, E. J., Neill, E., & Rossell, S. L. (2015). Assessing the Relationship between Semantic Processing and Thought Disorder Symptoms in Schizophrenia. *Journal of the International Neuropsychological Society*, *21*(8), 629–638. https://doi.org/10.1017/S1355617715000648

Tanskanen, P., Ridler, K., Graham, K. M., Haapea, M., Veijola, J. M., Jääskeläinen, E., Miettunen, J., Jones, P. B., Bullmore, E. T., & Isohanni, M. K. (2010). Morphometric Brain abnormalities in schizophrenia in a population-based sample: Relationship to

duration of illness. *Schizophrenia Bulletin*, *36*(4), 766–777. https://doi.org/10.1093/schbul/sbn141

Tarchi, L., Damiani, S., la Torraca Vittori, P., Marini, S., Nazzicari, N., Castellini, G., Pisano, T., Politi, P., & Ricca, V. (2021). The colors of our brain: an integrated approach for dimensionality reduction and explainability in fMRI through color coding (i-ECO). *Brain Imaging and Behavior*. https://doi.org/10.1007/s11682-021-00584-8

Tăuţan, A.-M., Rossi, A. C., de Francisco, R., & Ionescu, B. (2020). Dimensionality reduction for EEG-based sleep stage detection: Comparison of autoencoders, principal component analysis and factor analysis. *Biomedical Engineering / Biomedizinische Technik*, *66*(2), 125–136. https://doi.org/10.1515/bmt-2020-0139

Thoma, R. J., Monnig, M., Hanlon, F. M., Miller, G. A., Petropoulos, H., Mayer, A. R., Yeo, R., Euler, M., Lysne, P., Moses, S. N., & Cañive, J. M. (2009). Hippocampus volume and episodic memory in schizophrenia. *Journal of the International Neuropsychological Society*, *15*(2), 182–195. https://doi.org/10.1017/S1355617709090225

Toda, M., & Abi-Dargham, A. (2007). Dopamine hypothesis of schizophrenia: Making sense of it all. *Current Psychiatry Reports*, *9*(4), 329–336. https://doi.org/10.1007/s11920-007-0041-7

Uno, Y., & Coyle, J. T. (2019). Glutamate hypothesis in schizophrenia. *Psychiatry and Clinical Neurosciences*, *73*(5), 204–215. https://doi.org/10.1111/pcn.12823

Van den Oord, E. J., Rujescu, D., Robles, J. R., Giegling, I., Birrell, C., Bukszár, J., ... & Muglia, P. (2006). Factor structure and external validity of the PANSS revisited. *Schizophrenia research*, *82*(2-3), 213-223. https://doi.org/10.1016/j.schres.2005.09.002

Van Der Maaten, L., Postma, E., & Van den Herik, J. (2009). Dimensionality reduction: a comparative. *J Mach Learn Res*, *10*(66-71), 13.

Vapnik, V., & Chervonenkis, A. Y. (1964). A class of algorithms for pattern recognition learning. *Avtomat. iTelemekh*, *25*(6), 937-945.

Venkataraman, A., Whitford, T. J., Westin, C. F., Golland, P., & Kubicki, M. (2012). Whole brain resting state functional connectivity abnormalities in schizophrenia. *Schizophrenia Research*, *139*(1–3), 7–12. https://doi.org/10.1016/j.schres.2012.04.021

Vita, A., de Peri, L., Deste, G., & Sacchetti, E. (2012). Progressive loss of cortical gray matter in schizophrenia: A meta-analysis and meta-regression of longitudinal MRI studies. *Translational Psychiatry* ù,2. https://doi.org/10.1038/tp.2012.116

Ward, L. M. (2013). The thalamus: Gateway to the mind. In *Wiley Interdisciplinary Reviews: Cognitive Science*, 4(6), 609–622. https://doi.org/10.1002/wcs.1256

Wang, Q. (2012). Individualized medicine, health medicine, and constitutional theory in Chinese medicine. *Frontiers of Medicine*, *6*(1), 1-7.

Wallwork, R. S., Fortgang, R., Hashimoto, R., Weinberger, D. R., & Dickinson, D. (2012). Searching for a consensus five-factor model of the Positive and Negative Syndrome Scale

for schizophrenia. *Schizophrenia research*, *137*(1-3), 246–250. https://doi.org/10.1016/j.schres.2012.01.031

Werner, S., Malaspina, D., & Rabinowitz, J. (2007). Socioeconomic status at birth is associated with risk of schizophrenia: Population-based multilevel study. *Schizophrenia Bulletin*, *33*(6), 1373–1378. https://doi.org/10.1093/schbul/sbm032

Weiden P. J. (2002). Why did John Nash stop his medication?. *Journal of psychiatric practice*, *8*(6), 386–392. https://doi.org/10.1097/00131746-200211000-00010

Wechsler, D., & Psychological Corporation. (1997). *WAIS-III: Administration and scoring manual : Wechsler Adult Intelligence Scale*. San Antonio, Tex: Psychological Corporation.

Westerhuis, J. A., Hoefsloot, H. C., Smit, S., Vis, D. J., Smilde, A. K., van Velzen, E. J., ... & van Dorsten, F. A. (2008). Assessment of PLSDA cross validation. *Metabolomics*, *4*(1), 81-89. https://doi.org/10.1007/s11306-007-0099-6

White, T., Magnotta, V. A., Bockholt, H. J., Williams, S., Wallace, S., Ehrlich, S., Mueller, B. A., Ho, B. C., Jung, R. E., Clark, V. P., Lauriello, J., Bustillo, J. R., Schulz, S. C., Gollub, R. L., Andreasen, N. C., Calhoun, V. D., & Lim, K. O. (2011). Global white matter abnormalities in schizophrenia: A multisite diffusion tensor imaging study. *Schizophrenia Bulletin*, *37*(1), 222–232. https://doi.org/10.1093/schbul/sbp088

Wickham, H. (2016). ggplot2: Elegant graphics for data analysis () [Computer software]. The Comprehensive R Archive Network. Available from https://ggplot2.tidyverse.org

Williams, B., Onsman, A., Brown, T., Andrys Onsman, P., & Ted Brown, P. (2010). Exploratory factor analysis: A five-step guide for novices. *Journal of Emergency Primary Health Care (JEPHC)*, *8*, 2010–990399.

Winterburn, J. L., Voineskos, A. N., Devenyi, G. A., Plitman, E., de la Fuente-Sandoval, C., Bhagwat, N., Graff-Guerrero, A., Knight, J., & Chakravarty, M. M. (2019). Can we accurately classify schizophrenia patients from healthy controls using magnetic resonance imaging and machine learning? A multi-method and multi-dataset study. *Schizophrenia Research*, *214*, 3–10. https://doi.org/10.1016/j.schres.2017.11.038

Wismüller, A., & Vosoughi, M. A. (2021). Classification of Schizophrenia from Functional MRI Using Large-scale Extended Granger Causality. http://arxiv.org/abs/2101.10471

Wolfers, T., Doan, N. T., Kaufmann, T., Alnæs, D., Moberget, T., Agartz, I., Buitelaar, J. K., Ueland, T., Melle, I., Franke, B., Andreassen, O. A., Beckmann, C. F., Westlye, L. T., & Marquand, A. F. (2018). Mapping the Heterogeneous Phenotype of Schizophrenia and Bipolar Disorder Using Normative Models. *JAMA Psychiatry*, *75*(11), 1146–1155. https://doi.org/10.1001/jamapsychiatry.2018.2467

Woodward, T. S., Menon, M., and Whitman, J. C. (2007). Source monitoring biases and auditory hallucinations. Cogn. Neuropsychiatry 12, 477–494. doi: 10.1080/13546800701307198

World Health Organization (2019). International Statistical Classification of Diseases and Related Health Problems (11th ed.). https://icd.who.int/

Wright, I. C., Psych, M. R. C., Rabe-Hesketh, S., Woodruff, P. W. R., David, A. S., Psych, F. R. C., Murray, R. M., & Bullmore, E. T. (2000). Meta-Analysis of Regional Brain Volumes in Schizophrenia. In *Am J Psychiatry* (Vol. 157).

Robin, X., Turck, N., Hainard, A., Tiberti, N., Lisacek, F., Sanchez, J., & Müller, M. (2011). pROC: an open-source package for R and S+ to analyze and compare ROC curves () [Computer software]. The Comprehensive R Archive Network. doi: 10.1186/1471-2105-12-77

Yan, W., Shao, H., & Wang, X. (2004). Soft sensing modeling based on support vector machine and Bayesian model selection. *Computers and Chemical Engineering*, *28*(8), 1489–1498. https://doi.org/10.1016/j.compchemeng.2003.11.004

Yang, H., Liu, J., Sui, J., Pearlson, G., & Calhoun, V. D. (2010). A hybrid machine learning method for fusing fMRI and genetic data: Combining both improves classification of schizophrenia. *Frontiers in Human Neuroscience*, *4*. https://doi.org/10.3389/fnhum.2010.00192

Yoshihara, Y., Sugihara, G., Matsumoto, H., Suckling, J., Nishimura, K., Toyoda, T., ... & Takei, N. (2008). Voxel-based structural magnetic resonance imaging (MRI) study of patients with early onset schizophrenia. *Annals of General Psychiatry*, *7*(1), 1-11.

Yushkevich, P., Dubb, A., Xie, Z., Gur, R., Gur, R., & Gee, J. (2005). Regional structural characterization of the brain of schizophrenia Patients1. *Academic Radiology*, *12*(10), 1250-1261. https://doi.org/10.1016/j.acra.2005.06.014

Zanetti, M. V., Schaufelberger, M. S., Doshi, J., Ou, Y., Ferreira, L. K., Menezes, P. R., ... & Busatto, G. F. (2013). Neuroanatomical pattern classification in a population-based sample of first-episode schizophrenia. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *43*, 116-125. https://doi.org/10.1016/j.pnpbp.2012.12.005

Zilles, K., Schleicher, A., Langemann, C., Amunts, K., Morosan, P., Palomero-Gallagher, N., ... & Roland, P. E. (1997). Quantitative analysis of sulci in the human cerebral cortex: development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture. *Human brain mapping*, *5*(4), 218-221.

Zhou, L., & Wang, H. (2012). Loan default prediction on large imbalanced data using random forests. TELKOMNIKA *Indonesian Journal of Electrical Engineering*, 10(6), 1519-1525. http://dx.doi.org/10.11591/telkomnika.v10i6.1323

Zhao, C., Zhu, J., Liu, X., Pu, C., Lai, Y., Chen, L., Yu, X., & Hong, N. (2018). Structural and functional brain abnormalities in schizophrenia: A cross-sectional study at different stages of the disease. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *83*, 27–32. https://doi.org/10.1016/j.pnpbp.2017.12.017

Zhao, Y. J., Lin, L., Teng, M., Khoo, A. L., Soh, L. B., Furukawa, T. A., Baldessarini, R. J., Lim, B. P., & Sim, K. (2016). Long-term antipsychotic treatment in schizophrenia: systematic review and network meta-analysis of randomised controlled trials. *BJPsych Open*, *2*(1), 59–66. https://doi.org/10.1192/bjpo.bp.115.002576

Zou, K. H., Tuncali, K., & Silverman, S. G. (2003). Correlation and simple linear regression. *Radiology*. 227(3), 617–622. https://doi.org/10.1148/radiol.2273011499