A Metagenomic Hybrid Classifier for Paediatric Inflammatory Bowel Disease

Benjamin Wingfield[1], Sonya Coleman[1]
[1]Intelligent Systems Research Centre, Magee Campus, United Kingdom
Email: wingfield-b@email.ulster.ac.uk, sa.coleman@ulster.ac.uk

T.M. McGinnity[1,2]
[2]School of Science & Technology, Nottingham Trent University, Nottingham, United Kingdom
Email: martin.mcginnity@ntu.ac.uk

Anthony J Bjourson[3]
[3]Northern Ireland Centre for Stratified Medicine, Derry/Londonderry, United Kingdom
Email: aj.bjourson@ulster.ac.uk

***Abstract*—Inflammatory bowel disease (IBD) is a group of inflammatory diseases of the human colon and small intestine. IBD symptoms are non-specific; diagnosis can be delayed because an invasive colonoscopy is required for confirmation. Delayed diagnosis is linked to poor growth in children. Imbalances in the human intestinal microbiome - the community of microorganisms that reside in the human gut - are thought to contribute to the development of IBD. Work done to date in classifying host health statuses from patterns in human microbiomes with supervised learning algorithms has focused on modelling *what* is present in the gut (i.e. a bacterial census) with the random forest algorithm. Metagenomic shotgun sequencing is required to understand what is *occurring* in the gut (i.e. gene functions) and is often cost prohibitive for hundreds of samples. However, gene functions can be predicted with the Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PiCRUSt) software package, which could represent a valuable source of new features. In this paper we investigate feature relevance across the feature set with the Boruta algorithm. We find that the majority of relevant features are from the predicted metagenome. Support vector machines (SVM) and multilayer perceptrons (MLP) are rarely used with microbiomic datasets but offer several theoretical advantages. To determine if the new features and alternative algorithms are appropriate, we experiment with a range of machine learning and computational intelligence algorithms. With the best performing algorithms we also implement a conditional multiple classifier system that can identify IBD presence, IBD subtype, and IBD activity from a non-invasive stool sample.**

I. Introduction

Inflammatory bowel disease (IBD) is a group of disorders that cause persistent inflammation of the gut. IBD caused 53,000 deaths worldwide in 2013 and its prevalence is increasing [1]. IBD symptomatology is generally nonspecific and diagnosis is usually confirmed via invasive colonoscopy, with consequent delays. Delayed paediatric IBD diagnosis reduces growth significantly and is linked to poor treatment outcomes [2]. In [3], a bacterial census of the intestinal microbiome, defined below, was combined with a machine learning approach in order to classify IBD presence and IBD subtype. This approach analysed the bacterial groups present in the gut of IBD and control subjects; measuring what is *occurring* in the gut (i.e. abundance of gene functions) could provide a superior feature set for a supervised learning algorithm.

The human intestinal microbiome consists of the symbiotic, pathogenic, and commensal organisms that reside in the small intestine and colon. Investigating the intestinal microbiome requires microbial DNA to be taken directly from an environmental sample (i.e. stool in order to study the gut). This analysis is known as metagenomics. Once a sample is collected, DNA is isolated and sequenced. Marker gene surveys can create a bacterial census with traditional Sanger sequencing. 16S ribosomal RNA (rRNA) is commonly used as a marker gene when investigating the intestinal microbiome because it is a cost-effective protocol and only bacteria can be identified, negating contamination from the host.

Metagenomic shotgun sequencing is required to determine what is *occurring* (i.e. gene functions). In shotgun sequencing DNA molecules are broken into many pieces. These pieces are sequenced in a highly parallelised manner in a process known as next generation sequencing. The strands are then recombined *in silico* to form a continuous sequence for downstream analysis. Metagenomic shotgun sequencing of dozens or hundreds of samples is often cost prohibitive. An algorithm that can infer gene functions from a bacterial census via a reference database has been developed, dubbed Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PiCRUSt) [4] and will be used in this work.

Supervised learning algorithms have been used to identify host health statuses from marker gene surveys of human microbiomes for a number of conditions including IBD, irritable bowel syndrome, and pregnancy [3], [5], [6]. Random forests have been the sole learning algorithm used to date. Although, according to the no free lunch theorem [7], there will be no single best model for all metagenomic datasets, algorithms such as support vector machines (SVM) and multilayer perceptrons (MLP) have found success in DNA microarray analysis but are rarely applied to metagenomic classification problems. An extension of the no free lunch theorem holds that multiple classifier systems are superior to a single "computational point of view", and can result in better performance [8].

This paper presents a novel conditional multiple classifier system (hybrid classifier) that is able to identify IBD presence, IBD subtype, and IBD activity from a stool sample in a non-invasive manner with the aim of reducing time to diagnosis and reducing the need for invasive colonoscopy. To date inferred metagenomes have not been used as a source of features for

disease classification. Additionally, multiple classifier systems have not been applied to the classification of IBD previously.

II. Background and related research

Industrialised western nations have the highest IBD incidence and prevalence - approximately 261,000 people suffer from IBD in the United Kingdom - and this has increased significantly worldwide since the start of the 20th century [1]. IBD symptoms include abdominal pain, weight loss, and diarrhoea. In severe cases surgical intervention is required and the inflamed parts of the gastrointestinal tract are removed. IBD is a complex disease with uncertain aetiology [9]. IBD has two major subtypes: Ulcerative colitis - the effects of which are limited to the gut - and Crohn's disease, which can affect the entire gastrointestinal tract. IBD is usually episodic and severe inflammation is considered to be active IBD. IBD can enter remission during periods in which limited or no symptoms occur. IBD diagnosis is slow in children because IBD has non-specific symptoms; Colonoscopy is a specialised procedure and IBD symptoms are required before colonoscopy will be used for confirmation. Thus further development of non-invasive tests for IBD would be valuable.

A non-invasive test for paediatric IBD has been developed by mapping the intestinal microbiome [3]. These tests use a supervised learning approach to identify IBD presence and IBD subtypes. Mapping the intestinal microbiome to classify health status has to date been limited to analysing the relative abundances of bacterial taxonomic groups (i.e. *what* is present in the gut). This approach has been the most commonly used algorithm for classification of health status from patterns in human microbiomes [3], [5], [6]. An extension of this approach is to analyse the diversity and abundance of bacteria in a modular fashion instead of investigating biological species. Bacterial species in a metagenome can differ significantly between subjects [10]. A consequence of this is that relevant features or trained models can suffer from poor generalisation. Random forests have found widespread use classifying samples into microbiome phenotypes due to the excellent performance of the algorithm.

Other valuable algorithms that have been used less frequently in metagenomic classification problems include SVMs and MLPs. SVM classifiers are robust to a large number of irrelevant features and high feature-to-sample ratios, and use regularisation techniques to avoid overfitting [11]. MLPs and deep learning have only recently been applied to metagenomic classification, and show good results [12]. Due to the reliance on subjective diagnostic criteria for diagnosing IBD activity, learning a model for IBD activity could represent a non-linear problem: hence different SVM kernels and MLPs were experimented with. Combining multiple classifier systems into a hybrid system has not been done to date in the field of metagenomics.

In summary the literature has focused on classifying host health statuses from surveys of bacterial groups present in the intestinal microbiome with the use of random forest ensemble learning algorithms. A proposed extrapolation of the no free lunch theorem is that no single computational model solves all problems, demonstrating the advantage of hybrid models [8]. The aim of this work is two-fold: i) to determine the relevance (via the Boruta algorithm) of what is *occurring* in the intestinal microbiome of IBD subjects by generating a predicted metagenome with PiCRUSt; ii) to test a variety of supervised learning algorithms with the novel PiCRUSt-derived features and build a hybrid classifier using the best performing features and algorithms.

III. Data preprocessing

16S ribosomal DNA was sourced from a publically available dataset [3]. In this dataset DNA was isolated from faecal samples of 158 children (control n=37, IBD n=*122*). The sequenced DNA was processed to remove poor quality sequences, technical artefacts, and chimeric sequences before being uploaded to the Sequence Read Archive. In this dataset a sample is defined as all DNA sequences identified by the 16S marker gene survey per faecal sample. The DNA sequences identify hundreds of different bacterial groups per sample.

A. Obtaining a vector representation of bacterial presence and abundance

The DNA sequences were mapped to vector representations in order to input them into supervised learning classification algorithms. An Operational Taxonomic Unit (OTU) approach was used to generate these vector representations from DNA sequences (see figure 1). Bacterial taxonomic groups were identified and clustered from the similarity of the DNA sequences present per sample. DNA sequences that are 97% similar are commonly binned into the same OTU. The Quantitative Insights Into Microbial Ecology (QIIME) software package [13] was used to generate OTU tables (a community data matrix), with the open-reference subsampled OTU picking algorithm. The OTU table recorded how many times an identified OTU occurred for each sample. OTU abundances were scaled to be in the range $[0, 1]$.

B. Obtaining a vector representation of functional genetic content

The Phylogenetic Investigation of Communities by Reconstruction of Unobserved States (PiCRUSt) algorithm [4] was used to infer functional content from the marker gene survey. Marker gene surveys only identify specific genes and can only identify taxonomic and phylogenetic features. Metagenomic shotgun sequencing is usually required to investigate metabolic and functional features. Deep metagenomic sequencing of hundreds of samples remains cost prohibitive. PiCRUSt infers genetic content from bacterial phylogenies via comparison to a database of reference genomes. The presence and abundance of gene functions present per sample were binned into categories and abundances were scaled to be in the range [0, 1].

C. Obtaining vector representations of subject clinical data

Subject clinical history was converted from categorical variables to indicator variables (e.g. has the patient been prescribed immunosuppressant drugs). Immunosuppressant drugs

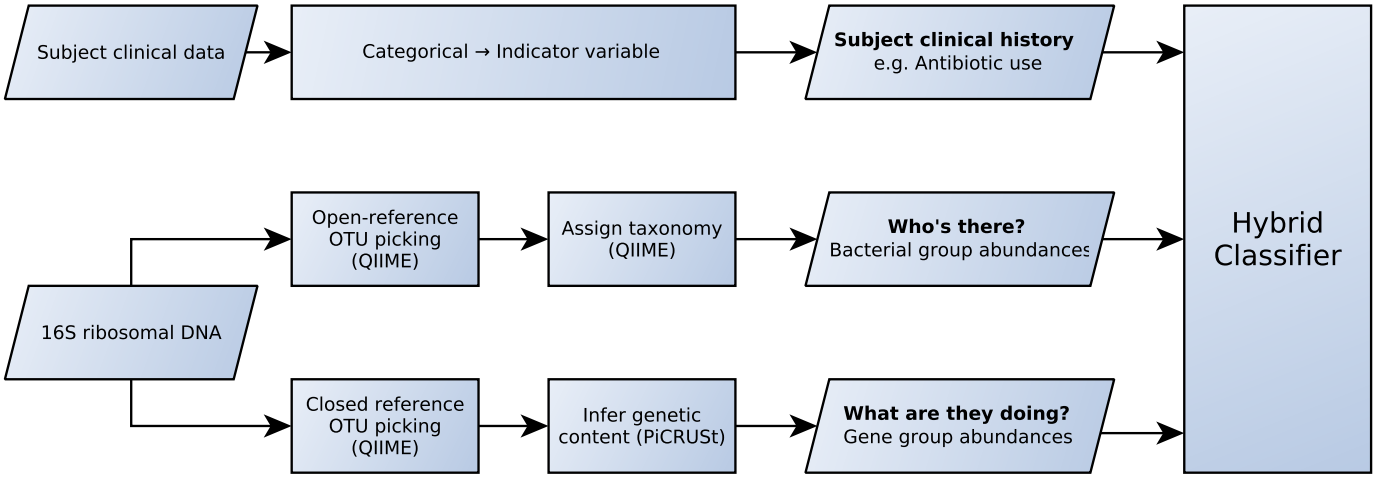


Fig. 1. Feature engineering pipeline. Subject clinical data is converted from categorical variables (e.g. family history of IBD) to artificial indicator variables. 16S ribosomal DNA was sourced from a publically available dataset (NCBI BioProjectID 82109 [3]), and consisted of the stool of 158 children (symptomatic controls n=37, IBD n=122) Operational Taxonomic Unit (OTU) picking clusters bacteria into groups from the similarity of their DNA. Typically if two DNA samples are at least 97% similar they can be considered to be in the same OTU.

and antibiotics [14] have been shown to cause large long term changes to microbiomes across the human body and it is thus essential to record their application. Other clinical data includes subject ethnicity and family history of IBD.

IV. Feature selection and feature relevance

Feature selection was applied in stage III with a MLP because high dimensional learning with traditional artificial neural networks is difficult [15]. 3-fold cross-validated SVM recursive feature elimination (SVM-RFE [16]) was used to automatically identify the optimal number of features. SVMs show good performance on high dimensional classification problems. SVM-RFE repeatedly eliminates the features least important (measured by SVM feature weights) to classification performance until the optimum is reached.

Boruta - an all-relevant feature selection algorithm based on a random forest wrapper and widely used in metagenomics [17] - was chosen to assess which features were relevant for each stage of the hybrid classifier (see Table I). By identifying every relevant feature in a classification problem the underlying mechanisms can be explored.

V. Multiple classifier system classification

Multiple classifier systems (hybrid intelligent systems) have many advantages: combined classifiers can outperform the best individual classifiers, multiple classifiers are more likely to find an optimal model, and multiple classifiers can be efficiently implemented in a multi-threaded environment in a parallel manner. The topology of multiple classifier systems is usually parallel or conditional. In parallel topology each classifier has identical inputs, and the final decision is made from the combined outputs of each classifier. In conditional topology, classifiers are used in a serial manner. Input is only passed to the classifier next in the sequence if some condition is met. Hybrid classifiers have found use in highly dimensional medical data sets including the classification of Alzheimer's disease from structural fMRI data. In this paper we used a serial classifier system. By returning a reduced set of classes at each stage of the serial classifier a complex problem can be iteratively decomposed into a series of simpler problems that are easier to classify.

A. Topology

The topology of the serial multiple classifier system (see Figure 2) was designed so that a complex problem (thorough IBD diagnosis) could be reduced to a set of simpler, but clinically important, problems. The presence of IBD (i.e. IBD or control) is important to guide treatment options. Some IBD treatments are contraindicated for subjects with conditions that can be misdiagnosed as IBD (i.e. prescribing immunosuppressant drugs for amoebic dysentery). The subtype of IBD and current IBD activity is important to guide the treatment course (e.g. severe Crohn's disease may require surgical intervention).

B. Classifier choice

The performance of SVMs and random forests was compared on the first two stages. These algorithms were picked because random forests have found widespread use in metagenomic classification problems, and SVMs have been shown to be on average superior to random forests on DNA microarray datasets [18]. DNA microarray data sets share many properties with metagenomic datasets such as high feature-to-sample ratios and large feature counts. SVMs were found to have superior performance to random forests for the first two stages.

The last class (IBD activity) is based on a subjective criteria in which various attributes are assigned a series of ratings (in the range $[0, 10]$). The attributes are summed and a diagnosis is given on the basis of this (e.g. using the Paediatric Crohn's Disease Activity Index [19]). A MLP was the best performing classifier for this stage.

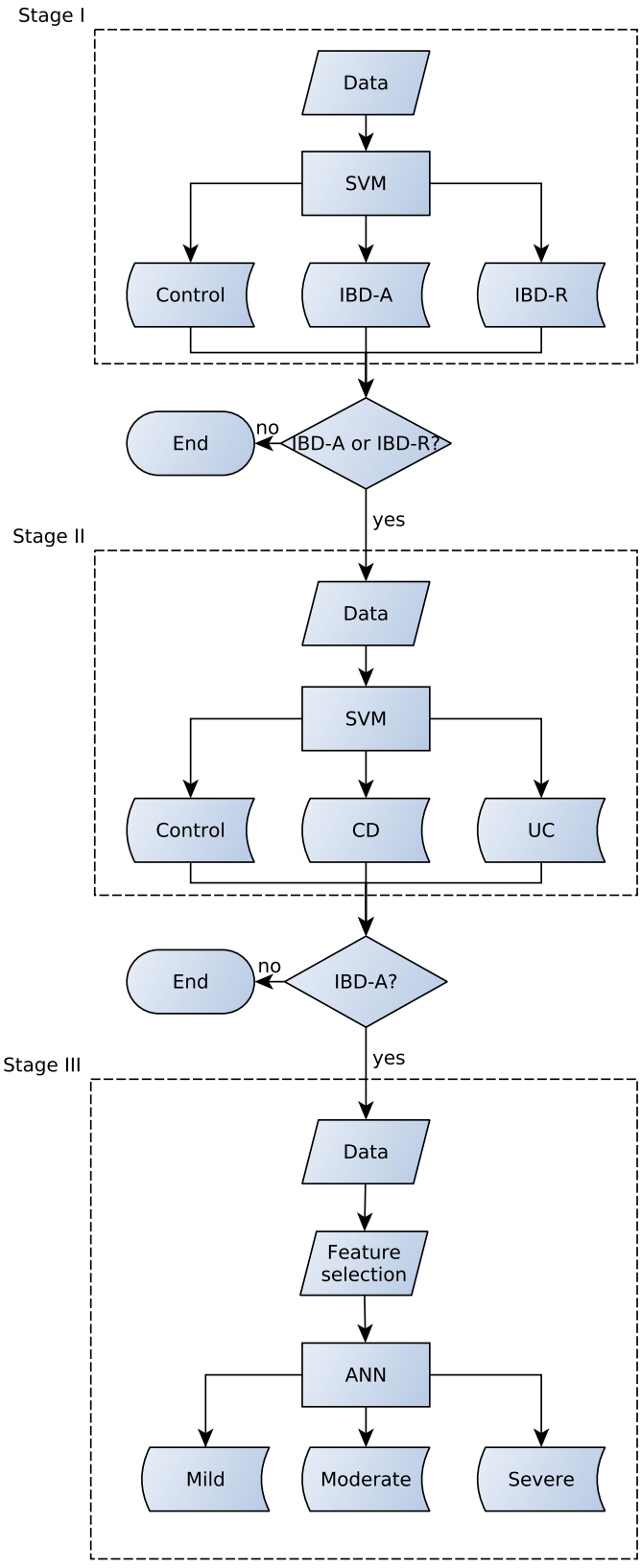


Fig. 2. Conditional multiple classifier system topology.

VI. Results

A. Relevance of novel feature set

The novel feature set of predicted gene functions was the most relevant type of feature across all three stages of the hybrid classifier (see Table I). To date no other metagenomic classifiers of host health status have used predicted metagenomes as a feature. However, classification of microbial communities with predicted metagenomes has occurred with good results [20]. The Boruta algorithm is an all-relevant feature selection algorithm that has found widespread use in metagenomics [5], [6]. Boruta considers a feature to be relevant if it is better for classification than a copy of the feature with a different order of values (dubbed shadows). Irrelevant features are iteratively pruned from the dataset until none are left. This is in contrast to traditional minimal-optimal feature selection methods, which seeks to find the smallest subset of features while optimising classifier performance. The authors of Boruta argue that minimal-optimal feature selection can lead to difficulty in detecting overfitting. The aim of Boruta is to understand the mechanisms of action that created the dataset.

Classification of disease state with a predicted metagenome has not been done to date. Adding the novel feature set tripled the number of features used from the original analysis of a bacterial census [3]. The novel feature set could cause the models to overfit on irrelevant features. Determining what features are relevant and if the relevant features are involved in mechanisms associated with IBD is an important step in validating the use of a predicted metagenome as a feature set for classification.

Carotenoid biosynthesis was a relevant feature in stages I and II. Carotenoids are a group of organic pigments synthesised by plants and bacteria, and are the pigments that produce attractive colours in plants. They are sourced mainly from fruit and vegetables and are antioxidants. The pathogenesis of IBD is thought to involve oxidative stress. In IBD patients antioxidants that circulate in blood plasma - including carotenoids - are present at significantly lower concentrations than controls [21]. This pattern is also found in this analysis but in this work the carotenoid biosynthesis is only measured from bacteria (plants do not have 16S rRNA). The intestinal microbiome synthesises a variety of important vitamins that are required by host metabolism such as vitamin B12. However, limited work has been done in assessing the role of the microbiota in carotenoid biosynthesis (e.g. vitamin A). Carotenoid synthesis by commensal bacteria could contribute to overall host health in previously undiscovered ways, and imbalances in the intestinal microbiome reduce the amount of carotenoid biosynthesis occurring.

Genes associated with bacterial infections were found to be relevant features in IBD subjects. There is evidence that conserved genes associated with *Vibrio cholerae* can be acquired by *Campylobacter concisus*, leading to the pathogenesis of IBD. *Vibrio cholerae* can increase the permeability of the intestine, triggering the onset and relapse of IBD. *Vibrio*

TABLE I
DISTRIBUTION OF RELEVANT FEATURES PER STAGE. FEATURE RELEVANCE CALCULATED WITH THE BORUTA ALGORITHM [17]. CLINICAL HISTORY INCLUDES SUBJECT PRESCRIPTION HISTORY (E.G. RECENT ANTIBIOTIC USE).

Stage	*Who's there? Bacterial abundance*	*What are they doing? Gene functions*	*Clinical*
IBD presence	27%	64%	9%
IBD subtype	34%	53%	13%
IBD activity	20%	80%	0%

TABLE II
CROSS VALIDATED AVERAGE PRECISION SCORE PER STAGE.

Stage	Average precision score	Support (classes balanced)
IBD presence	0.71	111
IBD subtype	0.65	111
IBD activity	0.61	45

cholerae genes are associated with the first stage of the hybrid classifier, which determines if IBD is present in its active form or remission [22].

B. Hybrid classifier performance

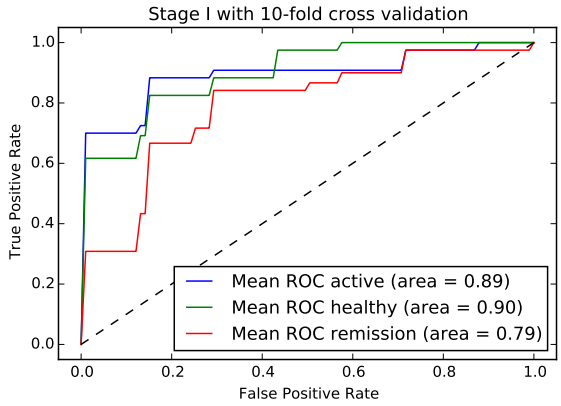
The first stage of the hybrid classifier showed good classification performance for the IBD remission class (see Figure 3a). The IBD active and control classes showed excellent performance. The second stage of the hybrid classifier had excellent performance for the failsafe control class and good performance for the Crohn's disease and ulcerative colitis (see Figure 3b). The third stage of the hybrid classifier had good performance for all classes (see Figure 3c).

Despite roughly tripling the amount of features when compared with the original analysis of a bacterial census (643 features were used including a bacterial census, predicted gene abundances, and clinical features in the final model) the SVMs used in the first two stages performed well. SVMs are insensitive to high feature-to-sample ratios. Random forests and SVMs performed poorly on the third stage of the hybrid classifier. Both could consistently identify mild and severe classes but were unable to classify moderate classes. A MLP showed good performance for all classes despite the nonlinearity of the data.

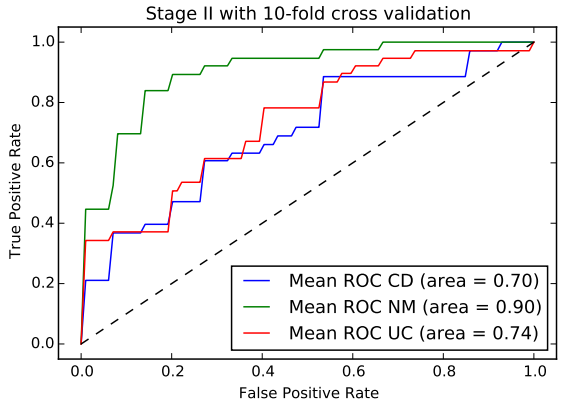
The average precision score of the last stage of the hybrid classifier shows the worst performance across all stages. This could be contributed to the lack of training data (see Table II) compared with the other two training stages. Additionally, classification of a subjective criteria (i.e. a class arising from a rating rather than a biological test) is a difficult problem.

C. Discussion

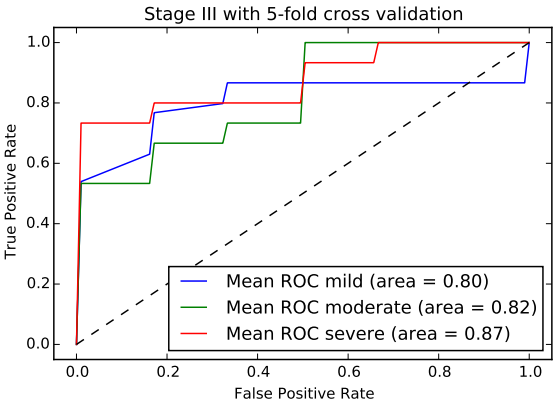
A method for the detection and stratification of IBD presence, IBD subtype, and IBD severity from a bacterial census of the intestinal microbiome, inferred genetic content of the bacterial population, and patient clinical history via a hybrid classifier was presented. An analysis of the relevant features



(a) Stage I: Classification of IBD presence with a support vector machine. IBD remission classification performance was good, IBD active and control classification performance was excellent.



(b) Stage II: Classification of IBD subtype with a support vector machine. Crohn's disease and ulcerative colitis classification performance was good, failsafe control class classification performance was excellent.



(c) Stage III: Classification of IBD severity with a multilayer perceptron. All classes showed good to excellent performance.

Fig. 3. Classification performance as measured by Area Under a Receiver Operating Characteristic Curve (ROC AUC) of the three-stage hybrid classifier. A ROC curve measures the ability of a test to correctly classify subjects and is a superior method of measuring performance to other metrics such as accuracy. A ROC AUC of 0.5 is a worthless test (equal to random chance, shown as the dotted diagonal line).

across all stages of the hybrid classifier showed that predicted genetic content was a valuable feature type, forming the majority of relevant features.

The performance of a classifier as measured by a Receiver Operating Characteristic (ROC) analysis is often measured via the area under the curve (AUC). The ROC AUC is often a better indicator of classifier performance than the misclassification rate or a loss matrix [23]. In [3] the sensitivity and specificity of the random forest classifier matched or surpassed alternative clinical methods (i.e. non-colonoscopy tests) for detecting IBD. The hybrid classifier presented in this work shows superior performance to the standalone random forest classifier.

An advantage of using a conditional multiple classifier system is its ability to maintain good performance for three different classification problems across nine classes. A MLP was the only algorithm capable of reliably classifying all classes for stage III, while SVMs showed superior performance for stages I and II. The three stages were designed to provide relevant information to a clinician which could guide treatment minimising the need for invasive colonoscopy. SVMs have been rarely used for metagenomic classification problems but show good performance in this case. Very little work has been done in applying MLPs to metagenomic classification problems. MLPs have shown value here for the classification of disease properties determined by subjective criteria. This is commonly done in many diseases with uncertain aetiologies, including depression.

The Boruta algorithm validated the use of predicted metagenomes as a novel feature set for the classification of IBD from marker gene surveys. Boruta revealed relevant features involved in biological mechanisms behind the pathogenesis of IBD. Significantly reduced abundance of antioxidant carotenoids in IBD subjects has been previously measured from blood plasma but not from the intestinal microbiome [21]. Carotenoids are typically sourced from fresh fruit and vegetables, as they provide plants with bright pigments. The role of the intestinal microbiome providing the human host with nutrients and vitamins has been well documented. No work to date has described the role of the intestinal microbiome in providing carotenoids to the host.

Genes associated with the lifecycle of pathogenic bacteria were also detected as relevant by the Boruta algorithm. Of note is the *Vibrio cholerae* lifecycle which is relevant for the first stage of the hybrid classifier that identifies IBD in remission, active IBD, and control classes. Evidence has been found that a combination of *Vibrio cholerae* and *Campylobacter concisus* is implicated in altering the permeability of the intestine, leading to IBD relapse into an active state [22]. This was previously missed by the bacterial census. Identifying bacterial species with a 16S marker gene survey is difficult due to technical limitations of the protocol. Species that are identified are typically present in very low abundances. This creates highly sparse feature vectors. Sparse data is challenging to learn from because it can increase the hypothesis space through which the learning algorithm must search.

VII. Conclusion and future work

In this work we present a novel combination of SVMs and MLPs in a conditional multiple classifier system that can discriminate the presence, subtype, and activity of paediatric IBD from a stool sample - avoiding the need for invasive colonoscopy - is presented in this work. SVMs and MLPs are rarely applied in metagenomic classification, but both have shown good performance. Work to date in associating human microbiomes with host health status has focused on analysing the bacterial groups present in the gut. Analysing what is *occurring* in a microbiome has required shotgun sequencing, which can be cost prohibitive for dozens or hundreds of metagenomic samples. The novel use of a predicted metagenome as a feature set, produced from a 16S marker gene survey with PiCRUSt, was validated with the Boruta algorithm. Metagenomic functions are a promising source of features for the classification of disease in a host. Features identified as relevant by Boruta have been independently implicated in the pathogenesis of IBD and IBD relapse. In each stage of the hybrid classifier the predicted metagenomes formed the majority of relevant features.

Future work will include validating the predicted metagenome in an independent cohort with metagenomic shotgun sequencing, including analysis of host genotype. Diseases with uncertain and complex aetiologies often have a genetic component. Discriminating between Crohn's disease and ulcerative colitis (stage II) was the most difficult stage for the hybrid classifier with the poorest performance. Ulcerative colitis only affects the intestine, while Crohn's disease affects the entire gastrointestinal tract - host genetic factors contribute to this difference. Fusing parallelised classifier decisions at each stage would further improve system performance and reduce uncertainty.

Acknowledgment

This work was completed under a PhD studentship supported by the Department of Education and Learning (DEL).

References

[1] N. A. Molodecky, S. Soon, D. M. Rabi, W. A. Ghali, M. Ferris, G. Chernoff, E. I. Benchimol, R. Panaccione, S. Ghosh, H. W. Barkema *et al.*, "Increasing incidence and prevalence of the inflammatory bowel diseases with time, based on systematic review," *Gastroenterology*, vol. 142, no. 1, pp. 46–54, 2012.

[2] C. Spray, G. Debelle, and M. Murphy, "Current diagnosis, management and morbidity in paediatric inflammatory bowel disease," *Acta Paediatrica*, vol. 90, no. 4, pp. 400–405, 2001.

[3] E. Papa, M. Docktor, C. Smillie, S. Weber, S. P. Preheim, D. Gevers, G. Giannoukos, D. Ciulla, D. Tabbaa, J. Ingram *et al.*, "Non-invasive mapping of the gastrointestinal microbiota identifies children with inflammatory bowel disease," *PloS one*, vol. 7, no. 6, p. e39242, 2012.

[4] M. G. Langille, J. Zaneveld, J. G. Caporaso, D. McDonald, D. Knights, J. A. Reyes, J. C. Clemente, D. E. Burkepile, R. L. V. Thurber, R. Knight *et al.*, "Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences," *Nature biotechnology*, vol. 31, no. 9, pp. 814–821, 2013.

[5] D. M. Saulnier, K. Riehle, T.-A. Mistretta, M.-A. Diaz, D. Mandal, S. Raza, E. M. Weidler, X. Qin, C. Coarfa, A. Milosavljevic *et al.*, "Gastrointestinal microbiome signatures of pediatric patients with irritable bowel syndrome," *Gastroenterology*, vol. 141, no. 5, pp. 1782–1791, 2011.

[6] K. Aagaard, K. Riehle, J. Ma, N. Segata, T.-A. Mistretta, C. Coarfa, S. Raza, S. Rosenbaum, I. Van den Veyver, A. Milosavljevic *et al.*, "A metagenomic approach to characterization of the vaginal microbiome signature in pregnancy," *PloS one*, vol. 7, no. 6, p. e36466, 2012.

[7] D. H. Wolpert and W. G. Macready, "No free lunch theorems for optimization," *Evolutionary Computation, IEEE Transactions on*, vol. 1, no. 1, pp. 67–82, 1997.

[8] M. Woźniak, M. Graña, and E. Corchado, "A survey of multiple classifier systems as hybrid systems," *Information Fusion*, vol. 16, pp. 3–17, 2014.

[9] S. B. Hanauer, "Inflammatory bowel disease: epidemiology, pathogenesis, and therapeutic opportunities," *Inflammatory bowel diseases*, vol. 12, no. 5, pp. S3–S9, 2006.

[10] M. Tong, X. Li, L. W. Parfrey, B. Roth, A. Ippoliti, B. Wei, J. Borneman, D. P. McGovern, D. N. Frank, E. Li *et al.*, "A modular organization of the human intestinal mucosal microbiota and its association with inflammatory bowel disease," *PloS one*, vol. 8, no. 11, p. e80702, 2013.

[11] B. Scholkopf and A. J. Smola, *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press, 2001.

[12] G. Ditzler, R. Polikar, and G. Rosen, "Multi-layer and recursive neural networks for metagenomic classification," *NanoBioscience, IEEE Transactions on*, vol. 14, no. 6, pp. 608–616, 2015.

[13] J. Kuczynski, J. Stombaugh, W. A. Walters, A. González, J. G. Caporaso, and R. Knight, "Using QIIME to analyze 16S rRNA gene sequences from microbial communities," *Current protocols in microbiology*, pp. 1E–5, 2012.

[14] E. Zaura, B. W. Brandt, M. J. T. de Mattos, M. J. Buijs, M. P. Caspers, M.-U. Rashid, A. Weintraub, C. E. Nord, A. Savell, Y. Hu *et al.*, "Same exposure but two radically different responses to antibiotics: Resilience of the salivary microbiome versus long-term microbial shifts in feces," *mBio*, vol. 6, no. 6, pp. e01 693–15, 2015.

[15] M. Verleysen *et al.*, "Learning high-dimensional data," *Nato Science Series Sub Series III Computer And Systems Sciences*, vol. 186, pp. 141–162, 2003.

[16] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Machine learning*, vol. 46, no. 1-3, pp. 389–422, 2002.

[17] M. B. Kursa, W. R. Rudnicki *et al.*, "Feature selection with the boruta package," 2010.

[18] A. Statnikov and C. F. Aliferis, "Are random forests better than support vector machines for microarray-based cancer classification?" in *AMIA annual symposium proceedings*, vol. 2007. American Medical Informatics Association, 2007, p. 686.

[19] J. S. Hyams, G. D. Ferry, F. S. Mandel, J. D. Gryboski, P. M. Kibort, B. S. Kirschner, A. M. Griffiths, A. J. Katz, R. J. Grand, J. T. Boyle *et al.*, "Development and validation of a pediatric crohn's disease activity index." *Journal of pediatric gastroenterology and nutrition*, vol. 12, no. 4, p. 449, 1991.

[20] Z. Xu, D. Malmer, M. G. Langille, S. F. Way, and R. Knight, "Which is more important for classifying microbial communities: who's there or what they can do?" *The ISME journal*, 2014.

[21] S. D'Odorico, R. Bortolan, R. Cardin, D. D'Inca', A. Martines, G. Ferronato, and A. Sturniolo, "Reduced plasma antioxidant concentrations and increased oxidative DNA damage in inflammatory bowel disease," *Scandinavian journal of gastroenterology*, vol. 36, no. 12, pp. 1289–1294, 2001.

[22] L. Zhang, H. Lee, M. C. Grimm, S. M. Riordan, A. S. Day, and D. A. Lemberg, "*Campylobacter concisus* and inflammatory bowel disease," *World journal of gastroenterology: WJG*, vol. 20, no. 5, p. 1259, 2014.

[23] T. J. Downey Jr, D. J. Meyer, R. K. Price, and E. L. Spitznagel, "Using the receiver operating characteristic to assess the performance of neural classifiers," in *Neural Networks, 1999. IJCNN'99. International Joint Conference on*, vol. 5. IEEE, 1999, pp. 3642–3646.