UNIVERSITY OF TARTU
Institute of Computer Science
Computer Science Curriculum

Samreen Mahak Hassan

Classification and Prediction of Business Incidents Using Deep Learning for Anomaly Detection

Masters's Thesis (30 ECTS)

Supervisor: Aleksandr Tavgen
Supervisor: Martin Kiilo
Supervisor: Raimundas Matulevičius, PhD

Tartu 2019

Classification and Prediction of Business Incidents Using Deep Learning for Anomaly Detection

Abstract: Companies today, uses a number of software systems to carry out various business activities. Such enterprise standard software solutions consist of a large number of components usually developed by different teams and/or different software vendors using various technologies. In such complex software systems, there can be various issues ranging from problems in the software itself to issues in network.

In order to measure the operational performance of applications and infrastructure as well as key performance indicators (KPIs) that evaluate the success of the organization, a lot of business metrics is collected. These metrics have certain data patterns which represent normal business behaviour. Anomalies are some unexpected changes within these data patterns such as degradation or sudden surge in business metrics values. Additionally, a small changes in software system configuration can cause unexpected behaviour in business flows. Version upgrades of different components can introduce compatibility problems. These problems could lead to change in normal behaviour of business metrics and cause anomalies. These anomalies if not resolved quickly results into business and financial losses. Therefore, it is necessary for businesses to take proactive steps to manage such business incidents before they can adversely affect it. This brings us to the need for an analytics platform which can analyze patterns of data streams, identify and differentiate normal behaviour of a business metric from anomalous behaviour and could generate notification.

The current anomaly detection and alert system in Playtech plc uses a simple anomaly detection technique that follows a rule based approach and it is observed that it is not efficient. Thus, a more robust, modular and efficient business incident/anomaly detection solution based on advanced machine learning techniques is needed that could work in conjugation with the current system. This thesis proposes, describes and evaluates a business incident/anomaly detection system based on deep learning approach that categorises and predicts the business incidents/anomalies using the available business metrics information.

Keywords:
Anomaly Detection, Business Incidents, Machine Learning, Deep Learning, Convolutional Neural Networks, Classification, Predictive Analytics, Playtech

CERCS: P170 Computer Science, numerical analysis, systems, control

Äriintsidentide klassifitseerimine ja prognoosimine kasutades süvaõpet anomaaliatuvastusel

Lühikokkuvõte: Tarkvarasüsteemid omavad tänapäeva äriettevõtetes elutähtsaid funktsioone ja nad on tihti ka äritegevuseks primaarse tähtsusega. Taolised süsteemid võivad koosneda väga suurest hulgast komponentidest, mis on arendatud erinevate meeskondade või ettevõtete poolt ning enamasti ka kasutades erinevaid tehnoloogiaid. Keerukate süsteemide korral võivad olla vead nii rakendustes kui ka võrgus.

Probleemid võivad ilmneda konfigureerimisel, mis võib põhjustada ootamatuid pöördeid ärivoos, samuti võivad versiooniuuendused tekitada kooskõlaprobleeme. See kõik võib põhjustada ärile maine- ja finantsilist kahju. Seetõttu on ärile vajalikud proaktiivsed sammud, et tulla toime äriintsidentidega enne nende ebasoodsat mõju teistele komponentidele. See toob kaasa vajaduse analüütilise platvormi järele, kus oleks võimalik eristada süsteemi normaalset käitumist anomaalsest meetrika alusel.

Playtech plc kasutab taoliseks automaatseks tuvastamiseks ja häirete tõstatamiseks tüüpilist anomaaliate tuvastamise lähenemist: reeglitel põhinevat tuvastamist. Playtech plc, tarkvarasüsteemides jälgitakse tuhandeid meetrikuid, alustades infrastruktuuri ja süsteemitarkvara ning lõpetades rakenduste ja ärimeetrikutega. Samas on tarkvara paigaldatud ja opereerib rohkem kui 40-s asukohas, igas neist erinevate lõppkasutajate ning ärimudelitest tulenevate erinevustega. Lisaks sellele, on tarkvara pidevas muutumises, nädalaste arendustsüklite tulemustena uuendatakse igal teisel nädalal komponente üle kõigi asukohtade ja paigalduste. Reeglitel põhinev lähenemine on piisavalt efektiivne tuvastamise kiiruse ja täpsuse osas, kuid nõuab palju inimressursse reeglite haldamise ja täppisseadistamise tõttu sellises muutuvas keskkonnas. Seetõttu nähti vajadust leida lahendus mis suutaks automaatselt kohaneda muutuvas keskkonnas ning erinevates tarkvara seadistustes ilma inimese pideva sekkumiseta. Antud töö eesmärk ongi masinõppel põhineva mudeli väljatöötamine ja treenimine, mis tuvastaks ja kategoriseeriks taolisi intsidente. Töö kirjeldab detailselt, kuidas kasutatakse anomaaliate tuvastamise ja süvaõppe tehnikaid täiendamaks olemasolevat lahendust intsidentide tuvastamisel ja klassifitseerimisel.

Võtmesõnad:
Anomaalia tuvastamine, äriintsidendid, masinõpe, süvaõpe, konvolutsioonilised närvivõrgud, klassifitseerimine, ennustav analüüs, Playtech

CERCS: P170 Arvutiteadus, arvanalüüs, süsteemid, juhtimine (automaatjuhtimisteooria)

IV. Licence

*Non-exclusive licence to reproduce thesis

I, **Samreen Mahak Hassan**,

1. herewith grant the University of Tartu a free permit (non-exclusive licence) to reproduce, for the purpose of preservation, including for the purpose of preservation in the DSpace digital archives until the expiry of the term of copyright,

   Classification and Prediction of Business Incidents Using Deep Learning for Anomaly Detection,

   supervised by Aleksandr Tavgen, Martin Kiilo and Raimundas Matulevičius.

   Publication of the thesis is not allowed.

2. I am aware of the fact that the author retains the right specified in p. 1.

3. This is to certify that granting the non-exclusive licence does not infringe other persons' intellectual property rights or rights arising from the personal data protection legislation.

Samreen Mahak Hassan
16/05/2019