



Research neurale netwerken en Software Cost Estimation

Reflectie

Bachelor in de 3ITF
keuzerichting AI

Naam

Academiejaar 2022-2023

Campus Geel, Kleinhoefstraat 4, BE-2440 Geel

1 INHOUD

In dit reflectie verslag van mijn stage ga ik even terugblikken op wat ik bereikt heb over de laatste 3 maanden. Ik heb een Erasmus stage afgelegd aan de Thomas Bata universiteit in Zlín, Tsjechië. Het onderwerp van mijn stage is een onderzoek naar neurale netwerken in correlatie met Software Cost Estimation (SCE) dit is het proces dat wordt geïnitieerd bij het begin van een nieuw software engineering project. In dit proces wordt er berekend hoe de werkkraft optimaal ingezet kan worden om een softwareproject te voltooien. SCE mag niet onderschat worden omdat het een directe samenhang heeft met het succes of falen van een project. Dit verslag bestaat uit 2 onderdelen, een inhoudelijke reflectie deze zal over wat ik bereikt op mijn stage en een persoonlijke reflectie waar ik de groei van mijn soft skills en kennis zal aanhalen.

2 INHOUDELIJKE REFLECTIE

Zoals ik eerder vermeld heb ging mijn stage over een onderzoek over Software Cost Estimation (SCE) in correlatie met neurale netwerken. In deze sectie ga uitleggen wat ik heb bereikt en wat dit oplevert voor de Thomas Bata universiteit. Mijn stage bestond uit 2 delen een review -en research paper.

2.1 Review Paper

In mijn Review Paper was het de bedoeling om een klein onderzoek te doen naar de bestaande methodologieën die al gebruikt zijn om SCE te doen, dit had twee functies: het onderwerp Software Cost estimation ontdekken/begrijpen en kijken wat er zoal in praktijk is gezet. Het kerndoel van de review paper was om te kijken waar de zogenaamde "Gap in Knowledge" (gat in de markt) zit. Ik ben erachter gekomen dat de meest voorkomende Artificiële Intelligentie techniek als het komt tot SCE, Machine Learning is. Dit vond ik zeer vreemd omdat in de laatste paar jaren Deep Learning krachtiger is geworden. De exacte reden waarom Deep Learning bijna nooit geïmplementeerd wordt ga ik uitleggen in de volgende sectie. Daarom had ik een voorstel gegeven om Deep Reinforcement Learning samen te implementeren met SCE. Mijn Review Paper zorgde dus ook als basis voor mijn Research Paper waar ik zelf ging experimenteren om diverse "gaps in knowledge" op te lossen.

2.2 Research Paper

In mijn Research Paper was het dus de bedoeling om zelf een onderzoek te gaan doen naar diverse methodologieën die nog niet gebruikt waren in het onderzoeksveld van Software Cost Estimation (SCE). Zoals vermeld in de vorige sectie is dit paper gebaseerd op de informatie uit de Review Paper. Ik heb eerst geprobeerd om SCE te doen aan de hand van Deep Reinforcement Learning (Voorstel Review Paper). Dit is een combinatie van Deep Learning en Machine Learning, hierdoor had ik snel de reden gevonden waarom Deep Learning niet geïmplementeerd wordt in Software cost estimation namelijk dat er niet genoeg data is. Dit heeft mij geleid naar het gebruik van synthetische data, deze wordt gegenereerd door een AI-model en behoudt de statische waarden van de originele data. Hierna ben ik deze synthetische data gaan evalueren op verschillende Machine Learning technieken. De resultaten waren zeer goed, dus daarom heb ik besloten om de modellen die het beste presteerden samen te gieten in een ensemble model. Deze heeft het voordeel om de sterktes van al deze algoritmes te combineren om daarna de finale predictie te maken. Het resultaat hiervan was zeer goed. In conclusie, het gebruik van synthetische data samen met een ensemble model geeft een accuraatheid van 100 percent gegeven mits de data in de 25 percent bereik ligt van de echte waarden.

2.3 Business case

2.3.1 Fakulta aplikované informatiky utb

Door middel van het onderzoek te publiceren krijgt de fakultiteit bekendheid, de reputatie van de school kan ook omhooggaan. Ze hebben natuurlijk ook nog de mogelijkheid om verder te gaan met het onderzoek omdat er nog enkele "gaps of knowledge" zitten die met verder onderzoek geëlimineerd kunnen worden.

2.3.2 Software engineering bedrijven | potentiële klant

Anderzijds kunnen Software engineering bedrijven ook een potentiële klant zijn, automatiseren van Software Cost Estimation bespaard veel tijd en kan ervoor zorgen dat een project vlotter verloopt.

2.4 Toekomstige Use-case

Zoals vermeld in sectie 2.3.1 is er nog altijd ruimte voor verder onderzoek naar aanleiding van de "gaps in knowledge" die ik ondervonden heb. Een voorbeeld hiervan is dat er geen gestandaardiseerde manier is om ensemble modellen te gebruiken voor Software Cost Estimation dit zorgt ervoor dat resultaten van verschillende onderzoeken verschillend kunnen zijn. Er zijn ook 3 verschillende methodes om ensemble methode te maken: bagging, boosting en stacking. Aan de andere kant is er ook nog en probleem met synthetische data omtrent privacy en data protectie. Als deze problemen opgelost zijn kunnen meer en meer bedrijven mijn methodologieën overwegen.

2.5 Tips voor opdrachtgever

De grootste tip die ik aan mijn supervisor zou geven is betere communicatie en meer mogelijkheden om meetings te houden waar ik mijn progressie kan laten zien. Over de 3 maanden heb ik hem maar 4 keer gezien en de communicatie verliep ruw. Aan de andere kant waren de meetings altijd wel zeer informatief omdat hij me goed kon verder helpen met zijn ervaring in dit onderzoeksveld.

3 PERSOONLIJKE REFLECTIE

In deze sectie ga ik het hebben over hoe ik de stage heb ervaren en hoe het mij heeft geholpen om een inzicht te krijgen van de professionele en businesswereld.

3.1 Ervaring

Ik heb deze stage positief ervaren, het was voor mij een heel grote stap om in het buitenland stage te gaan doen. Ik wist dat ik hierdoor op een persoonlijke en professionele manier ging groeien, meer hierover kan u lezen in de sectie 3.2 Groei. Het was een zeer interessante maar uitdagende stage. Ik heb nog nooit professionele papers geschreven met de intentie dat deze ook gepubliceerd zullen worden. Het onderzoek gedeelte was ook nieuw voor mij omdat de opdracht begint met een vage beschrijving van wat er verwacht wordt maar je moet voor de rest alles zelf uitzoeken naar aanleiding van wat jij denkt dat de beste manier is. Het zorgt ervoor dat er veel experimentatie plaatsvindt.

3.2 Competenties

Ik heb tijdens deze stage veel competenties opgedaan zoals formeel schrijven omdat ik dit nog nooit op deze schaal heb gedaan. Ik heb ook veel bijgeleerd over Software Cost Estimation en het belang ervan in de software engineering industrie. Ik ben er ook achter gekomen dat synthetische data zeer veel potentieel heeft in AI-projecten waar er weinig data beschikbaar is. Ook heb ik veel kennis opgedaan over diverse Machine Learning algoritmes zoals Linear Regression, Lasso en Support Vector Machines en hoe deze gecombineerd kunnen worden in een ensemble algoritme.

3.3 Groei

Ik heb tijdens deze stage een professionele en persoonlijke groei doorgemaakt, qua persoonlijke groei ben ik meer onafhankelijk geworden. Dit komt natuurlijk erdoor omdat ik alleen woonde in een vreemd land met een vreemde taal, hieraan heb ik mij goed aan kunnen aanpassen door middel van les te volgen in het Tsjechisch. Op professioneel vlak ben ik ook gegroeid qua skills, kennis en communicatie door een duik te nemen in het gebied van onderzoekers.

3.4 Problem solving

Problem solving was redelijk moeilijk tijdens het onderzoek omdat er weinig tot geen informatie was omdat ik een "Gaps in knowledge" proberen op te lossen. Daarom moest ik veel informatie opzoeken over de individuele onderdelen van het onderzoek.
