**Synoptic Project (End Point Assessment)**

**Module title:** Synoptic Project (EPA)  
**Module code:** COMP3932   
**Year:** 2022-2023 **Assignment:** Project summary – employer assessment/ sign-off

The synoptic project is a work-based project that broadly represents the skills, knowledge and behaviours in the [Digital and Technology Solutions Professional (Integrated Degree) standard (ST0119](https://www.instituteforapprenticeships.org/apprenticeship-standards/digital-and-technology-solutions-professional-integrated-degree-v1-1)). It should provide substantive evidence from a business-related project to demonstrate the application of skills and knowledge.

|  |  |
| --- | --- |
| **Apprentice name:** | Jake Kadir |
| **Apprentice SID:** | 201308103 |
| **Project title:** |  |
| **Occupational specialism:** *Please delete as appropriate* | Software Engineer/ Data Analyst |
| **Employer:** | PricewaterhouseCoopers |

|  |
| --- |
| **Please provide an overview of your synoptic project which is suitable for a non-specialist  (i.e. without using technical terms or jargon), and summarise the main achievements of your project.** *(no more than one A4 page)* |
| This study evaluates the use of natural language processing techniques to extract features from recipes, aiming to use the resulting features in content-based recommendation systems. A brief introduction is given to machine learning, deep learning, natural language processing and recommender systems.  Three approaches to feature extraction are given: using a pre-trained natural language model (BERT), training new natural language models (using both Word2Vec and fastText) and synthesising a dataset using BERT embeddings and classifiers. Two baseline techniques not using natural language processing (a text search and TF-IDF) are presented. A testing framework is outlined to evaluate these recommendations without existing recommendation labels.  A user study was conducted to collect data assessing the system's performance, and the results found that the performance of the Word2Vec implementation far exceeded the other techniques, suggesting there is significant merit in training a natural language model to learn embeddings that specifically capture the semantics of the recipes they represent.  A large body of ideas for future work is included. Suggestions include assessing the proposed techniques in different tasks, using additional sources of information in recommendation (e.g. images or text from non-English languages) and adapting the proposed techniques to use recipe properties to a greater extent. |

|  |
| --- |
| **Please explain how your project is applicable to the business environment.** *(250 words)* |
| The project investigates the process of intelligently recommending recipes. This could help employees find healthier meals, improving their well-being, or find recipes that use up leftover ingredients, helping them save food waste and money. In a broader sense, the research made into natural language processing and recommender systems is highly transferable into other projects. Many PwC projects use datasets of unstructured text, so investigating techniques of automatically extracting features from this data could result in useful, time-saving tools throughout the business. Recommender systems have many applications too and could be used to recommend content to employees (e.g. Digital Lab artefacts) or in developing products for clients. |

**Signatories**

|  |  |  |
| --- | --- | --- |
| **Apprentice:** | **Jake Kadir** | **Date: 28/04/2023** |
| **Employer:** |  | **Date:** |