Academic Background Introduction

An overview of the education and research experience of Han Tang for the application of ESR AIDD doctoral programs

9.9.2021 Han Tang

Educational & Research Background Overview

- BSc. In Applied Chemistry in Beijing University of Chemical Technology (BUCT) (2013 - 2017)
- Pre-master for MSc. In Computing (2017-2018) in Dublin Institute of Technology (2017-2018)
- MSc. In Computing (Data Science) in Technological University Dublin (2018-2020)
- Research Assistant in Technological University Dublin (2020-2021)

Undergraduate time in BUCT

- Modules for Fundamental Chemistry: Organic Chemistry, Inorganic Chemistry, Physical Chemistry, Structure Chemistry, Analytical Chemistry
- Modules for Fundamental Mathematics: Single Variable Calculus, Multi-variable Calculus, Linear Algebra, Probability & Statistics.
- My undergraduate thesis is in computational chemistry:
- Thesis Title: Study Factors Influencing the Layer Heights for Layered Double Hydroxides Using Computing Approaches

Graduate time in TU Dublin

- Pre-Master in Computing: Academic Writing, Object-Oriented Programming, Database Design.
- MSc. In Computing (Data Science): Data Mining, Machine Learning, Deep Learning, Probability & Statistical Inference, Data Visualisation.
- Thesis: A Comparison Study on State-of-the-art Minority Class Data Oversampling Techniques for Imbalanced Learning.

A Comparison Study on State-of-the-art Minority Class Data Oversampling Techniques for Imbalanced Learning

- Research on SOTA approaches for data re-sampling techniques for imbalanced learning.
- Review the Synthetic Minority Oversampling TEchnique (SMOTE) and its 11 extensions.
- Conduct experiments on the extensions on a large volume of imbalanced datasets. Evaluate the results and compare their performances, reached an inductive result.
- *For the high dimensional structure dataset, the extensions of SMOTE would not present a statistically significantly better result than the naive SMOTE method.*
- **Drawbacks**: Compare the performances are too straightforward and lack insights into the relation between dataset and the data re-sampling technique. If I designed the research object to be more specific, I could have potentially reached results of higher novelty.

Research Assistant in TU Dublin

- Cooperate with a tech startup AutoPlan to automate the planning application process for architects.
- Develop a method to find precedents for planning applications.
- The method is a Similarity Ranking Algorithm which ranks the similarities of planning applications in the dataset for a given planning application for recommendation (Recommender).
- The approach uses a linear function that includes features selected by expert architects to calculate the similarity scores.
- Developed a text classifier using FastText, which generate new features from the application texts.

Q & A