Transfer Learning of the Personality Scores in NLP HR Analytics

Author

Kuan Ying He

ucabkhe@ucl.ac.uk

Project Supervisors

Dr. Daniel Hulme

d.hulme@cs.ucl.ac.uk

Industry Partner - Oleeo

Dr. Ben Auffarth

ben.auffarth@oleeo.com

MSc Business Analytics with specialisation in Computer Science

Department of Computer Science.

University College London.

September 5, 2019

This report is submitted as part requirement for the MSc in Business Analytics (with specialisation in Computer Science) at University College London. It is substantially the result of my own work except where explicitly indicated in the text.

Abstract

In the job market, increasing numbers of people from overseas apply for jobs in the UK. Simultaneously, the number of job vacancies in the UK has been increasing. In this context, traditional recruitment processes may be unable to tackle the increasing numbers of job applications, leading to companies missing talented applicants. Therefore, for this research, NLP transfer learning will be applied to create a machine learning system to improve recruitment process.

In cooperation with Oleeo, a software company in London, the researcher is developing a machine learning solution for Oleeo's clients (companies) to improve the efficiency of their recruitment process. Specifically, this NLP solution is to predict job applicants'personalities based on the text of their CVs, which could be a next generation product from Oleeo. For this task, the researcher compares and contrasts different neural network structures, including the latest transfer learning pre-layers - Google XLNet. In this way, Oleeo's clients can achieve a more time- and cost-effective recruitment process.

Four experiments were conducted. Most significantly, the accuracy of the model used in experiment 3 is approximately RMSE 0.52, but with a correlation coefficient of 0.18. This means that the model is not completely ready for the market, due to the low correlation coefficient, but is nonetheless already capable of utilising NLP methods to predict candidates'personalities. Regarding the business value of the solution, it is expected to increase the efficiency of a client's recruitment processes by 30%. Moreover, it is expected that this service can increase Oleeo's business revenue by 15% in the next 3 years..

Project deliverables have formed the foundation for integration into the Oleeo product.

Acknowledgements

I want to express my gratitude to the following people who have made the work of this dissertation enjoyable and possible.

Dr. Ben Auffarth for project supervision and research guidance from Oleeo.
Dr. Daniel Hulme for research guidance from UCL Computer Science.
Richard Poulten for computing architecture, network setting support from Oleeo.
All the guidance, and support from Jack Hobson, head of technology at Oleeo
All the support from Stefan Sawh, Global Head of Account Management at Oleeo
All the guidance, and support from Charles Hipps, chairman at Oleeo
My classmate, Zhaoda Zhang, for statistical support from UCL.

Endorsements

Kuan Ying joined Oleeo Plc as an MSc thesis student. He was working on advanced text models and natural language processing. His particular focus was applying a model, XLNet, released by Google Brain in collaboration with Carnegie Mellon in mid-June this year, to predictive models from applicant resumes. He worked autonomously on a comparison involving a baseline model, adapting the XLNet model and running in different configurations, always coming up with new ideas for possible improvements.

He contributed to the knowledge in the data science team, and this will contribute to our modeling efforts in the future.

Ben Auffarth — Lead Data Scientist
Oleeo — 5-7 Bridgeworks The Crescent London SW19 8DR

Contents