The Future of Work - 12653

Author: Nian Yang Terence Tan   Supervisor: Dr. Michael A Osborne

November 16, 2022

1   Overview of the project

The world population is ageing over the next few decades (Gerland et al., 2014). The rising elderly to working age population ratio is increasing and will continue to do so (World Health Organization, 2022). This trend is known as an ageing population, and will strain the public and social services of many countries around the world (Wiener and Tilly, 2002). As one of the key social challenges facing the world for the next few decades, it would be interesting to examine how an ageing population will affect the economy, and in particular, the job market.

2   Key project objectives

In this project, we aim to examine the relationship between the age distribution within occupations and the degree of automation (both present and future) of those occupations. We might also look into any correlations with the skills/knowledge required for those occupations. This will all be done using a Bayesian non-parametric machine learning technique called Gaussian Process (Ghahramani, 2013).

The original goal was to extrapolate the changes in distribution of age groups within occupations over the next few years, but this was deemed unfeasible due to the limited amount of time-series data.

3   Progress to date

Over the summer, I have looked into Gaussian Processes and completed an online laboratory assignment designed by my supervisor. I hae also read past papers looking into similar areas of research.

Potential data sources have already been identified. The distribution of age groups within occupations are published by the US Bureau of Labour Statistics (BLS)[1], and go back to 2011. We can find the skills and knowledge associated with the various occupations on O*NET Online[2].

A GitHub repository for this project has also been created[3]; it contains codes, reports, etc.

4   Immediate tasks

Data wrangling work has already been started on the BLS data in order to standardise the data sets across the years. The differences are mainly due to changes to the Standard Occupational Classification (SOC) system. The SOC codes would allow us to map between the the BLS and the O*NET dataset.

---

[1] https://www.bls.gov
[2] https://www.onetonline.org
[3] https://github.com/terencetan-c/4YP-The-Future-of-Work

Since the changes to the SOC codes are quite complex, it was deemed unrealistic to automate the data wrangling. Instead, it would be done manually, which would still be manageable given that only about a hundred rows of data have been flagged for closer inspection. All the details of the data wrangling process can be found in an excel file on the GitHub repository.

5   Plan of work to the end of the project

Once the data wrangling is completed, figures and graphs will be produced from the standardised BLS data to give us a clearer picture of the data. We will then move on to selecting an appropriate Gaussian Process and train our model on the BLS data. If things go well, we may then look at the O*NET data and incorporate it into this project. While doing all of these, I will start working on the report and the presentation.

A high-level Gantt Chart was created to allow us to track our progress over time; the list view of the chart can be seen in Figure 1.
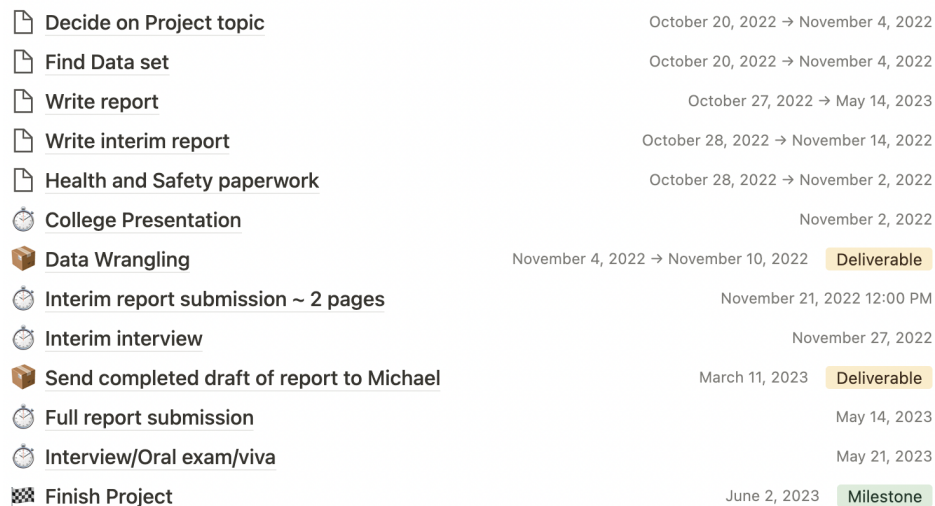

Decide on Project topic	October 20, 2022 → November 4, 2022
Find Data set	October 20, 2022 → November 4, 2022
Write report	October 27, 2022 → May 14, 2023
Write interim report	October 28, 2022 → November 14, 2022
Health and Safety paperwork	October 28, 2022 → November 2, 2022
College Presentation	November 2, 2022
Data Wrangling	November 4, 2022 → November 10, 2022 Deliverable
Interim report submission ~ 2 pages	November 21, 2022 12:00 PM
Interim interview	November 27, 2022
Send completed draft of report to Michael	March 11, 2023 Deliverable
Full report submission	May 14, 2023
Interview/Oral exam/viva	May 21, 2023
Finish Project	June 2, 2023 Milestone

Figure 1: Gantt Chart in List View

References

Gerland, Patrick et al. (2014). "World population stabilization unlikely this century". In: *Science* 346.6206, pp. 234–237.

Ghahramani, Zoubin (2013). "Bayesian non-parametrics and the probabilistic approach to modelling". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 371.1984, p. 20110553.

Wiener, Joshua M and Jane Tilly (2002). "Population ageing in the United States of America: implications for public programmes". In: *International journal of epidemiology* 31.4, pp. 776–781.

World Health Organization (Oct. 2022). *Ageing and Health*. URL: https://www.who.int/news-room/fact-sheets/detail/ageing-and-health.