

Find my NOC

CISC 372 – Project Proposal

Vedant Srinivasan Kartik Srinivasan | 17vsks@queensu.ca

Background

What is NOC and why is it important?

- National Occupational Classification is a system outlined by the Canadian government to classify jobs into different categories based on their duties and requirements. It is a way for jobseekers to find out what kind of jobs are in their sector and if their job duties are like the definitions prescribed.
- Apart from simplifying data collection for research on workforce and occupational trends, it is used as part of the criteria for new jobseekers and immigrants to complete work that is eligible for different government programs like Express Entry.

Problem Statement

- Imagine that you are a jobseeker, and you are looking for job postings online. You see many jobs, but you are not sure of which category the job is part of. You want to make sure that the job's duties are matching with the criteria. Of course, you could check online on the NOC directory but what if the job posting clearly told you what NOC it belongs to. Now, you know that this job is recognized through the NOC definition.

Solution

- A black box that takes in a job posting and gives out its NOC. This black box could be integrated in popular job sites to help jobseekers.
- This will also help the jobseekers find out the skills required for that NOC and related job titles.

Designing the Solution

Data

- The data could come from scraping publicly available job sites and it will consist of text in the form of the job posting.

Processing Data

- The text data must be pre-processed into a format that works best for our classification model. Rigorous cleaning procedures must take place such as stop-word removal, stemming, lemmatization, etc.
- We could use a vector embedding technique to convert the text into numerical format.

Machine Learning Model

- Given that the NOC is a hierarchical structure, I would use a technique that involves making a model for each level for the hierarchy. This would work better as some NOC codes are better represented than others. This means some models would give a higher accuracy than others.

References

<https://www.canada.ca/en/immigration-refugees-citizenship/services/immigrate-canada/express-entry/eligibility/find-national-occupation-code.html>
<https://noc.esdc.gc.ca/Home/AboutTheNoc/27f7a62adeb041b98171b81b774f6adb>