Resume Extraction - Group Project

Group name: Overloading

Buse Küçükçoban, buse_kucukcoban@hotmail.com, Turkey, Istanbul Technical University, NLP

Arda Cem Özmen, ardaozmen3@gmail.com, Turkey, Istanbul Bilgi University, NLP

Siddhartha Vanam, vanamsid@gmail.com, USA, University of North Carolina at Chapel Hill, NLP

Problem description:

Resumes contain surfeit information that is not relevant for the HR/authority, and they have to manually process the resumes to shortlist the promising candidates for them. And, thus making the shortlisting task a herculean task for HR. By making use of the NER(Named Entity Recognition) model of NLP this problem can be solved by finding and classifying the entities that are present in each resume into predefined classes such as person name, college name, academics information, relevant experiences, skill set etc.

Business understanding:

Classification of data using the NER model

Project lifecycle along with deadline:

Business understanding and planning project

Data analyzing, detecting missing, NA values etc.

Named Entity Recognition(NER)

Model Building & Training

Performance Evaluation & Reporting

Model Deployment

Model Inference

Data Intake Report

Name: NLP – Resume Extraction

Report date: 19-04-2021

Internship Batch: LISP01

Version: 1.0

Data intake by: Arda Özmen, Buse Küçükçoban, Siddhartha Vanam

Data intake reviewer: Data Glacier

Data storage location: https://gist.github.com/Rahulrky

Tabular data details:

Total number of observations	200
Total number of files	1
Total number of features	2
Base format of the file	.json
Size of the data	1.05 MB

Github Repo link: https://github.com/ardaozmen/Resume_Extraction