Resume Extraction - Group Project - Week 9

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.

Data understanding:

Classification of data using the NER model Dataset Attributes
{
'content': resume content or free text 'label': Annotated Tagged Entities

}

Data Cleaning and Transformations:

We classify data quality problems that are addressed by data cleaning and provide an overview of the main solution approaches. We have done classic nlp data cleanup jobs using regex and python language in data cleaning.

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