

Mining the US Technologist data

Abulitibu Tuguluke
DHI Group Inc.
6465 Greenwood Plaza Blvd
Greenwood Village, CO 80111
abtu8803@colorado.edu

ABSTRACT

What interesting question(s) were you seeking to answer? What is a brief summary of your results?

1 INTRODUCTION

Description of your question(s) Why they are important

2 RELATED WORK

Before this project, my personal interest are skill sets and knowledge graph for each skills. What I did was to use POS to set up skill key words and training on 1000 data set, so that to get the labeling for new data set working pretty well, afterward, I clustering multiple job description for certain job titles and found the common skill words for each job.

3 DATA SET

Where from Attribute features etc.

My approach for this project to apply a simple ML algorithm to a messy data, and not applying a complicated algorithm to a relatively nice data¹ Because personally, I just want to learning how to scrape and mining mutiple messy data and combine the knowledge.

3.1 EDA

4 MAIN TECHNIQUES APPLIED

Data clean/preprocess/etc. Data Warehouse/cube/etc. Classification/Clustering/etc.

4.1 Data scrapping

4.2 Data cleaning

4.3 Fuzzy matching

5 KEY RESULTS

What did you discover/learn?

5.1 BLS data set

6 APPLICATIONS

How can the knowledge gained be applied?

7 [OPTIONAL EXTRA CREDIT] VISUALIZATION

Visualization of results (and/or link to interactive site to view or link to a video demonstrating the interactive site)

¹I probably will do that in the Neural Net and Deep Learning class this fall.