Resume Writeup

Tell me About Yourself:

Hello, my name is Arhan Jain, an undergraduate from ceramic engineering. Since my childhood, I had been involved in a variety of activities along with academics as I loved to explore new areas. Similarly, in college, I tried to explore every opportunity in my college, along with managing my academics, I was actively involved in student activities which grown my interest in management and planning. There were times where data played a crucial role in planning and decision beside the tech world, which made me realize its importance. This started my journey towards data analytics and basic machine learning.

Internship:

Aim of the project was to develop an NLP model which could answer questions asked from an excel dataset directly without need of writing SQL query. To start this, my mentor first suggested few courses on NLP for me to complete to have a fair understanding about the concept (like Introduction to NLP in python-analytics Vidhya, Applied data science in python- Coursera). They provided me with a fair share of knowledge about how NLP helps computers to process human linguistics as computers only understand numbers. There were a few open-sourced models available on Github for me to implement like SQLNet, AllenNLP, but finally, I came across Tapas, which was an open-sourced model developed by Google. It used a training corpus from Microsoft(3.5GB file) which consisted of a question and their answers. Due to time constraints, I was not able to dig into the more profound understanding of the concept but gained a fair knowledge of how data was processed. (Data was extracted using NLP and turned into sequence then sequence to sequence encoding decoding was used to generate an answer for the asked question). The problem in the model was designed to handle directly coded data, so changes in the code was required in model libraries to take large data set imported from pandas. Finally, on a test set, we achieved 71% accuracy, which was correct prediction/total predictions.

pain point:

* SeqtoSeq encoding decoding
* Libraries used: keras,tensorflow, nltk
* LSTM

Zomato Project:

I started this project to get a better grasp over data visualization, which I learned from Applied Data Science with Python Specialization, how matplotlib and seaborn are used to plot different graphs.

About data:

The data had details of restaurant in a city including their complete address (even longitude and longitutde), type of cuisine served, cost, rating and other facilities available.

The data was processed enough already I had to only extract the data according to my requirement.

Insigths :

83 cities, 55,568 restaurant(major were north Indian ,then Chinese followed by fastfood and desserts). After comparing the average pricing for two, the average price in Tier 1 city was around **50% higher** than that in Tier 2 or Tier 3 cities **Dining chains** understood it and changed their prices accordingly.

Sentimental Analysis: this project aimed to predict whether a review on amazon for a cell phone is positive or negative.

First step was data preprocessing of dataset which had 2 columns review and rating. Using regex I removed most of the inconsistency in reviews then stemmed the data. That data was then used to generate bag of words which was converted to vector for model training.

Pain points:

AUC

Bigram

Tf-idf (could also have been used)

Strength: My greatest strength is that I am always planned and can manage stuff well. Like during my position as joint secretary of photography club there times where multiple events occurring at the same time in college. Once, our fest spardha was live and at the same time I was incharge of publicity of another fest which was after a week, so because of efficient planning and management we were able to complete both the tasks with reduced workforce.

I am very optimistic person. Give real life example when everyone thinks they have things will not turn out good , I always think that

Weakness: Most of the time I tend to take task given to other members who are under performing when we are near a deadline, this makes me to overwork and exhaust myself. Later I learned a lesson that like every finger in our hand is not same but all have a particular task. Similarly I started looking at everyone’s expertise and assigned them task accordingly to get the most out from everyone and efficiently.(real life example)

Walk me through your resume:

Well my resume sums up all of my college life and learning. As a curious and eager first yearite, I explored every thing that I could beacuase it was like w whole world infront of me ready to be explored. I participated in every freshers competiton and even won first place in two later i developed passion for photography then I joined the photography club enjoying my first year. In scond year I was elected as the joint secretary of it which helped me to develop a key skill which was planing and management. Along which I was the publicity coordinator of FMC weekend which taught me the importance of data……

All of this I managed along side balancing my academics, and also won Gymkhana Award

So to the field were we use to have play with data is data analytics and machine learning this started growth of professional life. **Intern and project**