Hi,

Thank you for giving me this opportunity to interview with JLL technologies.

Speaking about my experience, I have done master’s in data science from Illinois Institute of technology.

1. Daten solutions Inc.

During my Masters, I was interning with Daten Solutions Inc. based in Schamburg as a Data Engineer.

We primarily used snowflake, for data migration, where we used snowpipe for incremental load of the finance data.

Using snowsql we created the dimensions and the fact table for preparing a star schema.

For implementing the SCD for dimensions and the fact table, we used merge and streams concept of snowflake.

We tested the data by measuring the null values, count totals and deduplication in Dev, SIT, QA and Prod environments.

The CI/CD pipeline to consume incoming data was orchesterated using prefect library in python, using which we scheduled the runs and also monitored failing jobs.

After business validation, we created power bi dashboards to showcase the finance metrics.

Our client had an adhoc request of a creating a report of these metrics in MS Excel for a third-party consultant, so I extracted the dataset in CSV and then automated a report using Match, vloopkup and index functions in excel.

1. Labelmaster

We had a practicum course where we would work in industry for some company, so in my case I was working for Labelmaster.

It was a four-month project, where we were asked to develop a forecasting model for forecasting sales for different product departments. But before starting the machine learning part, we were given about 100 excel files which contained their historic sales data. We had to consume these files and upload it into snowflake. These files lie in the s3 bucket created by the company.

For that purpose, we created a stage in snowflake, in this stage we loaded the files from s3. We uploaded the stage files into a snowflake table, and post that we validated the results with the business team.

After validating the results with the business team, we started the predictive modeling process in Jupyter notebook. We had about 15 departments, we had to create a forecasting model for each department using about 200 exogenous variables. As we were working with the time series datasets, we used RNN neural network variety for forecasting the sales price accurately.

We managed to achieve the given benchmark of 8% error rate as compared to the actual sales value.

At the end we visualized our results using Tableau, where we showed month-on-month actual vs predicted values.

1. Cartesian Consulting

At cartesian consulting, as we were a consulting company, I got an opportunity to work with multiple clients. Starting with a grocery client, I was working for More grocery.

We had migrated their database to our system using Pentaho which was a migration tool.

So you have to select the tables of your interest and accordingly convert it to your database configuration.

For incremental data they used to send daily excel files through stftp, and I had created an automated script which would add the incremental data to bill header and bill details. So if there was an update in the master table, then we would expire the previous record and update the new records.

After setting up datawarehouse, we use to get campaign requirements from the more marketing team, hence for that purpose we had to apply the required machine learning model and extract the applicable customer base.

Further, we were also asked to create reports by our client for showcasing key KPI reports.

For that purpose, we have used macros in excel for creating the automated reports, or if sometime they want to see activity visually then we can create tableau dashboards of the same KPIs.

1. Greeksoft Technologies Pvt. Ltd.

Greeksoft is a trading firm,