Sandeep Kumar

(+91) 7060334397
skgrewal.95@gmail.com

EDUCATION

2014-2018

Indian Institute Of Technology, Roorkee

Bachelors of Technology

TECHNICAL STRENGTHS

Programming Languages: Python, PySpark, Sql.

Packages: Pandas, Numpy, Seaborn, Scikit-learn,Matplotlib, NLTK,plot.ly

Tools : Anaconda (Spyder, Jupyter notebook), databricks,Hdinsight, HDFS, Apache Kafka

Machine Learning: Classification, Regression, Time series prediction, NLP

Cloud Platform: Azure, AWS

Tools: GitHub, Jira , Assembla.

OS: Windows, Linux(Ubuntu), Debian.

WORK EXPERIENCE

4th July 2018 – Present

Associate Data Scientist, ATCS, Jaipur

Recommendation engine for repair of automobiles : This product is a recommendation engine for a luxury automobile. Implemented in python and sql on Azure cloud platform. Currently being used in over 30 countries by 3000 dealers.

- Historical repair data and master data for repairs is used to build recommendations.
- Used various statistical filters for data preprocessing and selection. Forest mapping is used to build operation chain.Similar parts are categorised into one based string similarity techniques.
- Technique Used : Collaborative filtering, Z- score, cosine similarity.

SEGMENTATION AND CHURN PREDICTION : This project is aimed at conversion of existing customers as potential new customers for lease and finance vehicles and finding the churn customer.

- Created multivariate classification attribution model and segmentation model using K-means clustering.
- Utilized high dimensional user service and financials dataset using PCA, LDA and Kernel approximation.Developed additive scoring method for marketing team.
- Deployed SGD, Logistic regression, random forest and XGBoost for classification model to boost sales by 30%

PREVENTIVE MAINTENANCE : We analyzed machine data to predict the anomaly in the machine.

- Classification models are used(logistic, LSTM and random forest) to categorise.
- Real time sensor data is utilized. Dashboard is created for monitoring the live status.

Prognosis of Reman factors *:* Demand Forecasting of value like return quantity for core parts to enhance inventory management and meet demand on time.

- Historical time series data is used. Data preprocessing and feature engineering to ensure stationary data.
- Visualisation plot, ACF Plots and statistical test like ADF test, P-valueTime are used to determine stationary data..
- Univariate Spline regression is used to determine the quantities for next 6 month with accuracy reaching upto 97% for some parts.

Support of PreSales activities, i.e. presentations, customer events, preparation of offers and efforts estimation

AGILE METHODOLOGY

- Good Knowledge of agile methodology
- Have done sprint planning, sprint retrospective

AWARDS/DISTINCTION

- Rewarded with '***Best of Best'-*** Award (Quarter 3,4, 2019) for exceptional performance, leadership & powerful solution **ATCS.**
- PM Scholarship Recipient

SKILLS

Data Analysis, Predictive modelling and analytics, Probability and Statistics, Project Management and Delivery