

DANISH NAWAZ

M-TECH (Industrial & Management Engineering)

Email: danisn@iitk.ac.in; Contact: +91-8582952206

LinkedIn: 

Github: 

ACADEMIC DETAILS

YEAR	DEGREE	INSTITUTE	CPI / %
Cont.	M.Tech (Industrial & Management Engineering)	Indian Institute Of Technology, Kanpur	9.76*(CPI)
2017	B.Tech. (Mechanical Engineering)	Government College of engineering, Berhampore	9.11(CPI)
2013	Class XII, West Bengal Board	M.G Rungta Academy, Kolkata	82.4 %
2011	Class X, West Bengal Board	Vidya Vikas High School,Kolkata	78.25 %

INTERNSHIP

Data Science Intern at Mphasis Nextlabs

Operation Process Recommendation System Using Machine Learning

(May'19- July'19)

Problem Statement: The Data was about farmer applying for agriculture relief funds schemes under the European Union Government. While application the system provides attribute level suggestion for individual applicants. It predicts the success probability, total time and number of steps to complete an application. It also Guide the Operation team to take **optimal path** for process flow and provides useful insights for success/failure of the journey. The project was divided into three sub-projects with solution approach as follows:

- Prediction system:** The farmers were divided into 5 clusters by attributes level **K-means clustering** followed by **Reference mapping** using average of normalized **L2(Euclidean) norm** and **Cosine similarity** to predict success probability and provide farmer's attribute level suggestion
- Recommendation System:** Performed **Feature Engineering** using Dynamic Functions coded from scratch, Applied **Random Forest** and **Support Vector Machine(SVM)** to find class level probability, which is used as a reward to recommend optimal path using **reinforcement learning** approach
- Statistical Analysis:** **Visualization** of failure journey using **directed flow graph (dfg)** under **PM4PY** library in python. Developed dynamic functions to find **point of divergence** and **reasons for failure**.

ACADEMIC PROJECTS

Data Mining	To Predict How Capable Each Applicant is of Repaying a Loan (Oct'18-Nov'18) -Applied K-Nearest Neighbour for data imputation, Principle Component Analysis(PCA) for data reduction, K-Fold Cross validation for data Re-sampling and removed multicollinearity -Applied models such as Support vector Machine (SVM) , Binary Logistic Regression and Decision tree . Support vector Machine (SVM) was the best suited model with area under ROC curve 0.756 and specificity 0.723
Statistical Modeling for Business Analytics	Analysis of the Factors Affecting Prices of Real Estates (Jan'19-Feb'19) -Carried out multivariate statistical regression analysis to study the factors influencing real estate prices -Calculated measure of fit, correlation matrix , performed EDA , heteroskedasticity test and checked for multicollinearity and omitted variable bias -Built log-linear , linear-log , log-log , non-linear models and finalized a model with adjusted R-squared value 0.79 and SER=0.203 Ethereum Price Prediction using Time Series Analysis (Mar'19-Apr'19) -Checked for Stationary , Seasonality , Trend using Dickey-Fuller test , EDA , ACF , PACF , Box-Ljung test for past 4 year data and predicted Ethereum closing price for 45 consecutive days -Built statistical models Exponential Smoothing , AR Model , ARIMA , ADL model and finalized model with RMSFE=37.22
Stochastic process	Credit Card Fraud Detection using Hidden Markov Models (Mar'19-Apr'19) -Used simulated data of a credit card user to train a Hidden Markov Model and estimated transition probabilities and emission probabilities using Forward-backward algorithm and sequentially predicted whether the upcoming transaction is fraud or not with recall 81% and F1 score 0.67
Marketing Research	Survey on Vodafone-idea Ltd to find market segmentation and reason for customer churning (Feb'19-Apr'19) -Designed cross sectional one shot case study dynamic survey form using Scaling techniques , Pretesting to control internal and external validity of Research Design. Data collected using online survey, Focus groups and Personal Interview -Conducted Exploratory , Descriptive Research in SPSS using primary data obtained by convenience random sampling -Analyzed sampled data using statistical test(Chi squared , One way ANOVA) to test our hypothesis - 32% of the customers marked customer service as the reason for churning while 28% marked price benefits by competitor. 48% switched to JIO just due to cost benefits . Customer retention is 81%
Decision Support Systems	Design a database incorporated System for 2 subprojects (Apr'19-May'19) -Designed user friendly and dynamic Decision Support System(DSS) to compare means of groups using one-way ANOVA method, leveraging HTML/CSS and PHP on server -Designed a Relational Database in SQL(MariaDB) for account opening form of a bank(HTML/CSS) and connected with Dashboard using PHP to store the data entered by user

COURSEWORK AND SKILLS

Relevant Courses	Data mining and knowledge discovery Probability and Statistics Machine Learning Introduction to computing-JAVA Operation Research Statistical Modeling for Business Analytics Advanced Statistical Modeling for Business Analytics Computer Aided Decision Support Systems Stochastic Processes Marketing Research
Technical Skills	Python (Numpy,Scipy,Panda,Matplotlib, Scikit-learn) R JAVA SQL HTML CSS PHP SPSS

POSITIONS OF RESPONSIBILITY

- Department Placement Coordinator at **Students Placement Office(SPO)**, IIT Kanpur (Mar' 19- Present)
- Serving as **Class Representative(CR)**, M.Tech Y18, IME Department, IIT Kanpur (Aug'18- Present)

ACHIEVEMENTS/AWARDS & CERTIFICATIONS

- Academic Excellence Award**, 2018, Recognition for outstanding academic performance across all departments (**awarded to top 7% students**)
- Working as a **freelancer at Chegg Inc**, provided **350+ solutions** for problems in **Probability and Statistics** with overall **rating of 95%** in 1 year
- Student of the year 2017** award for all round development in academic (**CPI 9.11**) and extra-curriculum activities from GCETTB
- SQL for Data Science**(4 weeks of hands on learning with assignments/projects),Grade obtained: 98.1%
- Sequence Model : A Certification in **NLP (Natural Language Processing)**