PARAS ARORA
MTech (Industrial & Management Engineering)

✉ aparas@iitk.ac.in | ☎ +91 8439039009
🔗 https://www.linkedin.com/in/parasarora72091aba/

ACADEMIC DETAILS

YEAR	QUALIFICATION	EDUCATIONAL INSTITUTION	PERCENTAGE
2018-20	M.Tech (Industrial & Management Engineering)	Indian Institute Of Technology, Kanpur	8.51* (CPI)
2013-17	B.Tech (Mechanical Engineering)	Harcourt Butler Technical University, Kanpur	72.60%
2012	Class XII (CBSE)	Krishna Public School, Meerut	88.8%
2010	Class X (CBSE)	Meerut Public School, Meerut	93.1%

upto 2nd semester

SUMMER INTERNSHIP

Data Science Intern at Core Compete Pvt. Ltd., Pune   *(May'19-July'19)*
- Objective of the project was to **Integrate Python and SAS Models in Model Manager using SAS Viya**
- Built **Churn prediction** models for telecom company in Python and SAS Model Studio
- It included Data pre-processing and Statistical analysis of variables and **K-fold cross validation** for resampling
- Classification Models used were **Logistic Regression, Random Forest** and **XGBoost**
- Python and SAS Models were registered in Model Manager and were used for scoring different datasets
- All the Models can be compared in Model Manager irrespective of the platform on which the model is built
- Suggested them to use **SWAT** for using SAS Viya through Python interface

ACADEMIC PROJECTS

Data Mining	**New York City Taxi fare prediction** *(Sept'18-Nov'18)* • The objective of the project was to predict taxi fare for New York city from the given dataset of 55 million rows • Steps included Data pre-processing , Data visualization, **Stratified Sampling**, **Feature Engineering** and Model building in R • Models used were **Linear Regression, decision tree** and **ensemble methods** like **Random forest** and **boosting** • Models were tuned and compared on basis of **RMSE** metric
Statistical Modeling for Business Analytics	**Predicting Income class using Logistic Regression using Adult data set from UCI Machine Learning Repository** *(Jan'19-Mar'19)* • The aim of the project was to predict whether a person's income is <50K or >=50K (target variable) based on given factors • Factors affecting income were "age", "education", "marital status", "gender", "income" etc • Data cleaning: Reduced the total no of factors in some columns and handled missing values and discrepancies • **Logit** and **Probit** models were used for classifying the income class. • Reported an **accuracy** of about **84.3%** ,**precision** of **61.9%** and a **recall** of **52.8%**, **AUC** of **ROC curve** was **0.88** **Predicting quality of Red wine using Statistical Regression Models** *(Feb-March'19)* • The data consisted of 1599 observations of 12 variables, Determined correlation matrix and checked for **Multicollinearity** • Carried out multivariate statistical regression analysis to study the factors influencing quality of red wine • **Breusch-Pagan** test showed heteroskedasticity, hence "**heteroskedastic robust errors**" were used • **Adjusted R2** without and with robust error was 0.3567 & 0.381 respectively • Statistically significant variables were alcohol, volatile acidity, density, chlorides and pH
Stochastic Process	**Credit Card Fraud Detection using Hidden Markov Models** *(Feb'19-Apr'19)* • Used simulated data of a credit card user to train a **Hidden Markov Model** and estimated transition and emission probabilities • Used **Forward-backward algorithm** and sequentially predicted whether the upcoming transaction is fraud or not • Reported **58% precision, 81% recall** and **F1-Score of 0.67**
Decision Support System	**Design a database incorporated System for 2 subprojects** *(Feb'19-Apr'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**	**Statistical Modeling for Business Analytics**| **Probability & Statistics** |Accounting and Finance |**Stochastic Processes and its Applications**	Advanced Corporate Finance	Security Analysis & Portfolio Management |Derivative Contracts |Mergers & Acquisition |Computer Aided Decision Support Systems| Operations Research for Management |**Introduction to Computing- JAVA**
Technical Skills	Python(NumPy, Pandas, seaborn, sklearn, keras,TensorFlow)| R|SQL| MS Office	PHP	HTML	SAS Model Studio

POSITIONS OF RESPONSIBILITY

- **Internship Coordinator** at Department of Industrial and Management Engineering, IIT Kanpur   *(Aug'18-Apr'19)*
- **Marketing Secretary** at Entrepreneurship Cell, IIT Kanpur   *(Feb'19-Present)*
- **Marketing Executive** at Association of Mechanical Engineers, H.B.T.I. Kanpur   *(Jul'14- May'16)*

ACHIEVEMENTS & CERTIFICATIONS

- Secured **99.5 percentile** in GATE Mechanical 2018.
- Secured **second** prize in Junkyard Wars, Mecharnival (Annual Techno-cultural fest)
- Secured 8970 rank in IIT-JEE 2013
- **Certified Business Analytics Professional** Course: Linear & Logistic Regression, Decision Tree, Segmentation, Market Basket Analytics
- **Machine Learning** (Breast cancer detection with SVM & KNN, Credit card fraud detection, Stock market **clustering**, Deep learning grid search, DNA classification, Text classification, Object recognition, PCA)