

MOHAMMAD AMIR

M. Tech. (Industrial and Management Engineering)

E-Mail – amirmd@iitk.ac.in | Contact - +91 7518387612

ACADEMIC DETAILS

Year	Qualification	Educational Institution	Board/University	Percentage
2018-20	M. Tech. (IME)	IIT Kanpur	IIT Kanpur	7.57/10*
2012-16	B. Tech. (ME)	AKGEC Ghaziabad	Dr. APJ AKTU Lucknow	70.92%
2009-11	Class XII	BRIC Harraiya, Basti	U.P. Board	84.40%
2008-09	Class X	MAICS Ganj, Kadirabad	U.P. Board	77.33%

*up to 2nd sem.

INTERNSHIP

Data Science Intern at Convergytics Solutions Pvt. Ltd.

May-July'19

- Objective was to **forecast** revenue for Tanishq at different granularities and **develop a frame work** for forecasting such data.
- Four years' worth of data was available with 29 attributes. Apart from this data for **Diwali and Akshay Tritiya holidays** was also available.
- There were lots of complexities in the data such as **large number of missing values, holiday and festival effects**. Data volume was low.
- Three scenarios** were considered for forecasting. These were "**Category wise**", "**Store wise**" and "**SKU wise**".
- ARIMA, TSLM, TBATS, NNETAR and LSTM (Long Short-Term Memory) deep learning model** were tried for forecasting.
- Out of these models **TSLM and LSTM** were performing best.
- KERAS and TensorFlow libraries** were used for **LSTM using Recurrent Neural Network**.

ACADEMIC PROJECTS

Data Mining	<p>Predict the three-month sales of 50 different items in 10 different stores</p> <ul style="list-style-type: none">Data consisted of daily sales of 50 different items in 10 stores from 2013 to 2017.Preliminary examination of time series for Store-1 and Item-1 showed that series was seasonal and non-stationary. Non-stationarity of series was confirmed by ADF and KPSS tests. Seasonality can also be seen from ACF and PACF plots.Three models were used for forecasting. These were ARIMA, FB-Prophet and NNETAR.ARIMA was not performing well due to high frequency of data. NNETAR model was overfitting the series.Prophet Model was the best model with SMAPE of 1.99.
Statistical Modelling for Business Analytics	<p>Predicting quality of RED WINE using Statistical Regression Models</p> <ul style="list-style-type: none">The data consisted of 1599 observations of 12 variables.Determined the correlation matrix and checked for multicollinearity.Carried out multiple linear regression with quality as dependent variable and all other as independent.It was found that four variables "citric.acid", "residual.sugar" and "density" and "fixed.acidity" were statistically insignificant hence we removed them from the model.Breusch-Pagan test showed that there is heteroskedasticity in the data.To take heteroskedasticity into account "heteroskedastic robust errors" were used for regression on significant variables.Adjusted R2 without robust error was 0.3567 and with robust error was 0.381 which is better than previous. <p>Predicting Income class using Logistic Regression</p> <ul style="list-style-type: none">The objective was to predict whether a person's income is <50K or >=50K based on factors such as "age", "gender" etc.Correlation plot showed no significant correlation between variables.Two models Logit and Probit were used for classifying the income class.Performance of both models were similar with accuracy of about 84.3% and precision of 61.9% and recall of 52.8%.AUC of ROC curve was 0.88 which shows that models' predictive power is good.

COURSE WORK AND SKILLS

Courses	Data Mining and Knowledge Discovery Probability and Statistics Statistical Modelling for Business Analytics Advanced Statistical Methods for Business Analytics Computer Aided Decision Support Systems Operations Research for Management Advanced Decision Models Introduction to Computing (JAVA) Introduction to Machine Learning (NPTEL)
Skills	R Python SQL PHP HTML MSOffice Excel

ONLINE LEARNING

- Introduction to Machine Learning (NPTEL)
- Practical Time Series Analysis (Course Era)
- Machine Learning by Building Projects (Eduonix)

POSITION OF RESPONSIBILITY

- Orientation Team Member (OTM) Counselling Service IITK.
- Discipline Committee Member UG.

AWARDS AND ACHIEVMENTS

- Secured all India score of 747 in GATE 2018.
- Finished in top 5% in online certification exam of NPTEL course **Introduction to Machine Learning**.