

# Mohammed Khan

## Data Scientist

Singapore

- Email me on Indeed: <http://www.indeed.com/r/Mohammed-Khan/56a77ba6bb5609fb>

Data Scientist with 3 years experience in Machine Learning and Statistical Methods and overall work experience of 6 years with a Masters Degree in Business Analytics from NUS, Singapore. Highly skilled in Data Analysis and Data Visualization with an excellent grasp of business operations.

## Work Experience

---

### Data Scientist

Charles and Keith - Singapore

June 2018 to July 2019

#### New Item Demand Prediction

Currently working on a Recurrent Neural Network (RNN) model for new fashion products using analogous product score from CNN models.

Modelled anELMNeural Network for demand prediction of Fashion shoes helping Business optimize supply chain and overall stock.

#### SKU Re-Order Prediction

Trained and optimized a RandomForest Classification model with 89% accuracy to predict item re-order, thus improving Sales by 15% for the next Quarter.

#### Singapore Sales Forecasting

Analyzed market data to understand Trend and Seasonality helping Business forecast Sales using different Time Series Regression techniques.

#### Bag's Classification Using OpenCV

Using OpenCV Blob detection, different Fashion Bags were classified using a Deep Learning ResNet classifier with 90% accuracy.

#### Sales and Marketing Analysis

Helped improve targeted marketing for specific Retail Stores by Cluster Analysis using K-Means clustering and analyzing sales KPI's.

Highlighted 'Red- Flags' to C-suite executives using intuitive Data Visualization.

### Data Scientist

NUS

January 2018 to April 2019

#### Kaggle Bike Sharing Demand Prediction

Trained an Ensemble model for the Bike Sharing Demand project from Kaggle.

Neural Networks, RandomForest and Gradient Boosting algorithm such as XGBoost were used for the Ensemble model.

#### Emotional Sentiment Multi-class Classification

Modelled various Classifiers to compare accuracy using Sci-kit Learn to Predict emotional sentiment from Brainwave readings.

RandomForest Classifier, Logistic Regression Classifier, XGBoost and Support Vector Machines classifiers were used for the comparison.

#### Boston House Price Prediction

Developed an XGBoost Regression model to predict the median house values with good RMSE using 5-fold Cross Validation approach. Used R and Python packages for Data Cleaning and Feature Engineering.

#### Car Park Availability - Big Data

Used Government data to show car park slot availability in Singapore for vehicle users.

Techniques- PySpark, HDFS, Hive, RandomForest Modelling and Tableau.

#### Recommendation System

Built Recommendation System with sequence mining on e-commerce click/buy data.

Sequential patterns were extracted using prefix projected pattern growth algorithm.

#### Time Series Regression - ARIMA

Developed an ARIMA Time Series model to forecast furniture sales using Decomposition techniques with good accuracy.

#### Predicting Parkinson Disease

Modelled a Linear Regressor to predict score of a patient affected by Parkinson's disease.

Used Principal Component Analysis for Dimensionality reduction.

### **Data Analyst - Infosys Ltd**

Infosys

January 2015 to May 2017

#### Credit Card Default Prediction

Developed Logistic Regression model to predict credit default based on customer's characteristics and past payment history.

Used R Programming for Data cleaning and Exploratory Data Analysis.

#### Credit Card Subscription Prediction:

Employed similar Logistic Regression technique to predict customer subscription for a second credit card based on customer data.

Used ROC curves for validations and Gains chart to visualize results.

#### Integrated Asset Classification

Analyzed and validated MorningStar and IDC vendor data for Asset Classification as per Business Rules for downstream applications.

Used MS Excel and GGPlot library from R for Reporting and simple visualization.

#### Customer Segmentation Analysis:

# Used K-Means Clustering technique to segment customers based on specific characteristics.

# Highlighted results to Portfolio Delivery Managers after cluster validations with the Business Analysts

### **QA Test Analyst**

Infosys Ltd

October 2012 to January 2015

3.5 years' experience in Functional, Automation and Database testing for Financial banking domain applications for a North American bank. Experienced in complete Software Testing Life Cycle with hands-on in Creating and Modifying Automation Scripts. Used HP ALM tool for Defect tracking.

## Education

---

### **Master's in Business Analytics**

National University of Singapore  
January 2018 to May 2019

### **Bachelor's in Engineering in Electronics and Instrumentation**

Anna University - India  
October 2008 to April 2012

## Skills

---

R (3 years), Python Programming (3 years), Machine Learning (3 years), Sql (4 years), Scikit Learn (3 years), Tableau (2 years)

## Links

---

<http://linkedin.com/in/mohammedismailkhan23>