

Junaid Butt

Profile

Data Scientist with multiple years of experience in performing data analysis to inform decision making and solve complex business problems.

Experienced in advanced analysis methods such as Machine Learning, Time-Series Forecasting and Natural Language Processing methods with an aim to help businesses develop a competitive advantage through effective leveraging of their data.

Moreover, I create robust, company-wide data driven cultures through the building of synergistic cross functional relationships with a view to align working pipelines and develop strategic solutions to deliver commercial success.

Contact:

Tel: +44(0) 7943 236501
Email: jmbutt28@gmail.com

Experience

Data Scientist

IBM, London United Kingdom/ December 2019 - Present

Natural Language Programming (NLP) experience with Python and R. Techniques used include sentiment analysis, information retrieval, topic modelling and vector space models.

Experience with network analysis measuring centrality, testing for randomness and community detection on large scale social media data. Overlaid these methods with NLP to investigate how conversation topics correlate to network activity.

Employed unsupervised ML techniques to quantify levels of polarisation in online conversation, one of the first applied uses of the Watson Debater service. Invented novel ways of measuring Watson Debater's effectiveness for classification tasks.

Deployed ML methods for bot/troll identification on social networks. Went beyond traditional model evaluation metrics like AUC to those which were better suited to the problem.

Worked with the Watson Machine Learning (WML) product team to add capabilities for the R programming language so that users could deploy machine learning models via API calls. Deployed ML models as a web service for selected clients using R Plumber.

Data Scientist

Pusher, London United Kingdom/ January 2018 - November 2019

Utilised Machine Learning methods to generate a lead scoring model for customers to determine probability to become a paying customer.

Built an automated forecasting tool with Bayesian Time Series methods and productionised it via R-Shiny which was adopted by all Salespeople company wide.

Performed analysis to understand how customers used Pusher products. Worked cross functionally with product and engineering teams to generate actionable insights from data using R and SQL.

Integrated new data sources into the data warehouse hosted by Google Bigquery. Subsequently integrated into the primary BI (Looker) tool which enabled company wide access to verified product and marketing data.

Skills

Standard SQL/ PostgreSQL/Oracle/ DB2

R & R Shiny

Python, Pandas, NumPy

Machine Learning

Hypothesis Testing (A/B)

GCP BigQuery

Git/ version control

Network Analysis

Natural Language Processing

R Markdown/ Latex

Self-starter and cross-functional

Looker and Tableau

Education

Birkbeck College
MSc Applied Statistics (2018)

University of Greenwich
BA Business Studies (2014)