DANIEL MUTUA

kithongo.mutua@gmail.com
www.linkedin.com/in/daniel-kithongo-mutua
+254-715-370-803
https://kitmutua.github.io/

>> DATA SCIENCE | ANALYTICS

MOTIVATION

My passion lies in applying Data Science & Advanced Analytics to tackle business obstacles innovatively and systematically. I consistently seek ways to enhance value for the team, business, and end-users, while perpetually expanding my knowledge.

SKILLS & TOOLS

Programming: Python (Base, Pandas, Numpy, Matplotlib, Scikit-Learn, Keras), SQL

Machine Learning: Linear Regression, Logistic Regression, Decision Trees, Random Forest, KNN, k-means, PCA, Association Rule Learning, Causal Impact Analysis

Other: Deep Learning, Statistics, Github, Data Visualisation, MS Office, Tableau, Jupyter Notebook

EXPERIENCE

Operations Management Team Lead - Invisible Technologies Inc.
DEC 2021 - PRESENT

- Implemented tailored data-driven strategies for a wide range of clients, resulting in a 40% increase in customer retention and a 15% boost in average order value.
- Directed the adoption of data-driven decision-making across the organization, resulting in a 50% reduction in operational costs and a 10% increase in overall revenue.
- Spearheaded a data-driven approach to decision-making, implementing predictive analytics models that reduced inventory costs by 15% and improved demand forecasting accuracy by 20%. costs by 15% and improved demand forecasting accuracy by 20%.
- Conducted in-depth consultations with clients to understand their unique data needs, leading to customized solutions that improved data insights by 50% and increased client satisfaction by 30%.

PROJECTS

Deep Learning Powered Image Search Engine

- Built a Deep Learning Image Search Engine that will help customers find similar products to the ones they want. The pre-trained VGG16 network was implemented and a Global Average Pooling Layer was added at the end of the VGG16 architecture for the output to be a single vector of numeric information rather than many arrays. The model was tested using two images and plotted the corresponding eight similar images and their cosine similarity scores.

Quantifying Sales Uplift With Causal Impact Analysis

- Used Causal Impact Analysis to analyze & understand the sales uplift of customers who joined the new campaign. There was a 41.1% uplift in sales for those customers who joined the campaign, over and above what we believe they would have spent, had the club not been in existence. It would be interesting to look at this pool of customers (both those who did and did not join the campaign) and investigate if there were any differences in sales in these periods the previous year - this would help the client understand if any of the uplift was because of the result of seasonality.

DANIEL MUTUA

kithongo.mutua@gmail.com
www.linkedin.com/in/daniel-kithongo-mutua
+254-715-370-803
https://kitmutua.github.io/

>> DATA SCIENCE | MACHINE LEARNING

Applying Association Rule Learning to quantify product relationships

- Used Association Rule Learning to analyze the transactional relationships & dependencies between products in the alcohol section of a grocery store. *Apriori* was used to examine & analyze the strength of the relationship between different products within the transactional data. The results could be used to identify the strongest or weakest relationships to guide the decisions regarding product layout, recommendations for customers, or promotions.

Leveraging Machine Learning for Enhanced Targeting Precision

- Applied Machine Learning to help a grocery retailer to reduce mailing costs, and improve Return On Investment. Four classification modeling approaches were tested: Logistic Regression, Decision Tree, Random Forest, and K Nearest Neighbors. Random Forest had the best scores among classification accuracy, precision, recall, and f1-score. The feature importance and permutation importance helped the client understand the key drivers behind delivery club signups.

Harnessing Machine Learning for Enhanced Retention Strategies

- Applied Machine Learning to help a grocery retailer predict the loyalty scores of the customers. Three regression modeling approaches were tested: Linear Regression, Decision Tree, and Random Forest. Random Forest had the highest predictive accuracy (95.5%). XGBoost and LightGBM could have been tested to check for the possibility of an increment in accuracy.

EDUCATION

Master of Economics
2017 - 2020 - Kenyatta University, KE

COURSES & CERTS

DSI Data Science Professional Certification

Actionable Learnings: Extracting & manipulating data using SQL. Application of statistical concepts such as hypothesis tests for measuring the effect of AB Tests. Utilising Github for version control, and collaboration. Using Python for data analysis, manipulation & visualization. Applying data preparation steps for ML including missing values, categorical variable encoding, outliers, feature scaling, feature selection & model validation. Applying Machine Learning algorithms for regression, classification, clustering, association rule learning, and causal impact analysis for measuring the impact of an event over time. Machine Learning pipelines to streamline the ML pre-processing & modeling phase. Deployment of an ML pipeline onto a live website using Streamlit. Using Tableau to create powerful Data Visualizations. Turning business problems into Data Science solutions.