KHATUNA KURDOVANIDZE

+447501892523
khatunamail@gmail.com
https://khatunakurdovanidze.github.io/
www.linkedin.com/in/khatunakurdovanidze

>> MSC STATISTICAL DATA SCIENCE | ML

MOTIVATION

*I am passionate about solving business problems using Data Science & Machine Learning. I **systematically** & creatively use my skillset to add tangible value to the team, the business, and the end-user. I am constantly learning, and always looking to improve.*

SKILLS & TOOLS

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

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

Other: Statistics (probability and Hypotheses Testing), Time Series, Github, Data Visualisation, MS Office, Jupyter Notebook, Tableau Dashboards and Stories

PROJECTS

Predicting Sales Using Advertising Data in R

Predict how much increase in sales will we get after spending money on different types of ads. I used Linear Regression for this task and according to linear regression formula and our data, if we spend $1000 on TV we increase sales by 19 units plus spending $1000 on radio gives us an additional 29 units plus the synergy effect 1.1 units, total 49 units. If we spend $2000 on both media, we get an increase in sales by 100 units.

Marketing Optimisation Using A/B Testing

The aim of this project was to help the business decide if they should spend money on the expensive mailing. We conducted the A/B test to assess the signup rate of two groups using the data from the previous marketing campaign of the company. Our results from the Chi-Square test for independence showed that it's not necessary to spend money on the expensive mailing and keep the simpler and cheaper one. This helps the company to save money while being effective during their future marketing campaigns.

Answering Business Questions with SQL

In this project I helped the business answer questions about their total sales, best customers, and the total sales for each product area. I used SQL(Structured Query Language) to accomplish this. This information can be used to support strategic decisions or to report latest results to stakeholders.

Targeting the Right Customers Using Machine Learning

How can a company reduce costs of the future campaigns by targeting only those customers who are more likely to sign up? I used Machine Learning algorithm called Random Forests to find out which customers have high probability to sign up for the Delivery Club. I have correctly classified all customers into two groups with low and high probability of signing up. Now the company can contact only the second group and reduce marketing costs.

KHATUNA KURDOVANIDZE

+447501892523
khatunamail@gmail.com
https://khatunakurdovanidze.github.io/
www.linkedin.com/in/khatunakurdovanidze

>> MSC STATISTICAL DATA SCIENCE | ML

COURSES & CERTS

Data Science Professional Certification (Data Science Infinity)

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 & visualisation. 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 & modelling phase. Deployment of a ML pipeline onto a live website using Flask & Heroku. Turning business problems into Data Science solutions.

Practical Time Series Analysis (Coursera)
Calculus - Mathematics for Data Science - Machine Learning (Udemy)
Linear Algebra for Data Science & Machine Learning (Udemy)

EDUCATION

MSc (Statistical Data Science), Merit
2020 - 2021 - University of Kent, UK

Actionable Learnings: Analyse the outcomes of the experiment held by an agricultural scientist who wanted to compare the effect of weed control on wheat yield to be able to detail conclusions for the nontechnical reader. I used ANOVA to analyse data and suggest the optimum treatment combination(cheap and effective). What the farmers could do in order to control weed the best and increase the yield is to use herbicide spray and combine it with nitrogen level 1(low) and potassium level 3(high). This way they could get almost five times more yield during the harvest time!

Using different data, I analysed data using Principal Component Analysis (PCA) and Factor Analysis (FA), decided on the number of components to keep and interpreted the results in terms of the loadings and the scores. Divided a different dataset into a training data set and a test data set, built a discriminant rule with LDA using the *training data*. Then I allocated emails either to *spam* or *not spam* in the *test data*. Used the confusion matrix to check the accuracy of the discriminant rule.

Forecasting the average number of goals for the English Premier League using data for the previous years. I used R programming language to choose the right model, conduct diagnostics and forecast. The data showed the average number of goals is decreasing with time, our model predicted that the average number of goals for the future four years will stay low.

BA (Psychology), 1st class
2010 - 2017 - Panteion University of Social and Political Sciences, Greece

EXPERIENCE

Other Administrative Roles
2004 - 2020