BIANKA MARKOVIC

biankamarkovich@gmail.com
www.linkedin.com/in/bianka-markovic/
https://github.com/biankamarkovic
https://biankamarkovic.github.io

DATA SCIENCE | MACHINE LEARNING

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

Tools: Excel, Tableau, Github, AWS (S3, Lambda, IAM, EC2, SageMaker, RDS, DynamoDB, Glue)

Math: Linear Algebra, Statistics (Hypothesis Testing, AB Testing, Central Limit Theorem, Distributions)

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

PROJECTS

"You Are What You Eat" Customer Segmentation

- Used k-means clustering on grocery transaction data to split out customers into distinct "shopper types" that could be used to better understand customers over time, and to more accurately target customers with relevant content & promotions

Compressing Feature Space For Classification using PCA

- Built a model to predict Ed Sheeran album buyers using PCA for data reduction and a Random Forest Classifier, achieving 93% accuracy to help target potential customers effectively.

Understanding Alcohol Product Relationships

- Applied Association Rule Learning using the Apriori algorithm to analyze product relationships in alcohol transactions, providing insights into product bundling and customer recommendations.

EDUCATION

BSc Geodesy & Geoinformatics

2016 - 2021 - University of Bonn, Germany

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 Streamlit. Using Tableau to create powerful Data Visualizations. Turning business problems into Data Science solutions.