Sistata Bagale

Sistata | sistata-bagale | shistatabagale14@gmail.com

Education

GEORGIA INSTITUTE OF TECHNOLOGY

MS in Analytics

August, 2023-current | Georgia, Atlanta, USA

KØBENHAVNS UNIVERSITET & TECHNISCHE UNIVERSTITÄT DRESDEN

Erasmus Mundus Joint Master's Degree in Sustainable Tropical Forestry

Aug 2020 - Oct 2022 | Copenhagen, Denmark | Dresden, Germany

Projects

CUSTOMER SEGMENTATION ANALYSIS

- ➔ Performed customer segmentation on a 2,240-record dataset using K-means and Spectral Clustering, identifying three distinct customer segments to drive targeted marketing strategies
- ➔ Engineered key features and applied Principal Component Analysis (PCA), reducing dimensionality while retaining 57.31% of variance, improving clustering accuracy.
- ➔ Optimized clustering models, achieving a Silhouette Score of 0.451 (K-means) and Dunn's Index of 0.017 (Spectral) for optimal segmentation.
- ➔ Derived actionable insights for tailored marketing strategies targeting low-income families, mid-income groups, and high-income premium customers.

COSMETIC PRICING ANALYSIS

- ➔ Developed a predictive model using machine learning techniques: Random Forest, LASSO, Stepwise regression) to analyze and forecast pricing dynamics.
- ➔ Preprocessed and cleansed 9,861 observations in R, utilizing binning and feature selection to enhance model performance and address multi-collinearity and data sparsity issues.
- ➔ Achieved superior predictive accuracy with the Random Forest model, demonstrating 29% reduction in Mean Squared Error

HOUSING PRICE ANALYSIS

- ➔ Developed a linear regression model to predict housing prices by identifying significant variables through Exploratory Data Analysis (EDA).
- ➔ Enhanced model performance by transforming input variables into second-degree polynomials, achieving an 18% improvement in adjusted R-squared value.

WEATHER PREDICTION

- ➔ Developed a time series forecasting model using Ridge regression to predict next-day maximum temperatures, reducing mean absolute error by 3.7% from 4.01°F to 3.86°F
- ➔ Processed and analyzed 19,393 rows of historical weather data by addressing 99.9% missing or invalid values through feature selection, forward filling, and rolling computations.
- ➔ Optimized model performance by engineering 18 new features, including rolling averages and percentage change metrics, resulting in a measurable improvement in prediction precision.

Skills

- Machine Learning
- Statistical Analysis
- Data Analysis & Visualization
- Research
- Report Writing

Tools/Platforms

- ➔ Python
- ➔ Numpy
- ➔ Pandas
- ➔ R
- ➔ Git
- ➔ Tableau
- ➔ SPSS
- ➔ ArcGIS

Relevant Courses

- Computational Data Analysis
- Regression Analysis
- Simulation
- Intro to Analytics Modeling
- Business Fundamentals for Analytics
- Simulation