Neslihan BABAARSLAN

Data Scientist

- Turkey
- +90 554 880 03 72
- n.babaarslan@gmail.com
- linkedin.com/in/neslihan-babaarslan/
- https://github.com/NeslihanBabaarslan
- https://www.kaggle.com/neslihanbabaarslan

PROFILE

3 years of experience as a **Data Scientist** with **Modelling Regression, Classification, Clustering** and creating, evaluating and managing data frames using large data framework. **5 years** professional career in **Data Analysis** with **MS Excel, Statistical Modeling, Predictive Modelling, Exploratory Data Analysis, Data Manipulation, Data Visualization.**

Academic research on **Healthcare** Datasets for the last **2 years** especially with **Deep Learning, Machine Learning, Data Processing, SQL, AWS**. Experience in performing **Supervised , Unsupervised ML and Deep Learning Algorithms** to train and test the big data sets. Also experience working on commonly used python libraries such as NumPy, Pandas, Seaborn, Matplotlib, Plotly, Scikit-Learn, Tensorflow. Experienced in working in **Agile environment. Strong analytical, creative thinking, problem solving, self-driven, highly motivated and pleasant personality.**

TECHNICAL SKILLS

- Python
- Keras
- Statistics
- Exploratory Data Analysis
- Seaborn
- SQL / MS SQL
- Git/Github
- AWS
- SciKit-Learn
- Tableau
- Machine Learning
- EC2-Sagemaker
- TensorFlow
- Agile/Jira
- Data Visualization
- Streamlit-Deployment

PROFESSIONAL EXPERIENCE

10/2020 – present	**Data Analyst & Data Scientist,** *EXIMMAN Manufacturing and eCommerce*

- Analyzing and visualizing keywords and product profitability,
- Product sales performance analysis that defines, measures and maintains using product metrics and **KPIs,**
- Preparing feasibility and investment reports for product launch,
- Data-driven product development using statistical analytics,
- **Predictive modelling** to increase and optimize customer service experiences,
- Understanding customer behavior and profile with **customer segment analysis,**
- Presenting result report and insights using **tableau,**
- **Modeling regression analysis** for sales estimates for Amazon US, EU marketplace.

02/2011 – present	**Labour Inspector, Data Analyst,** *The Ministry of Labor and Social Security*

- Evaluating the effectiveness of workplace inspectors by data-based analysis of annual inspection reports.
- Collecting data from workplaces, **creating a Healthcare dataframe,** analyzing using many methods and reporting in order to improve working condition.
- Developing and interpreting a statistical model (regression, classification, etc.) on the created healthcare dataframe.
- Providing insights into reasons for failure and success from **exploratory data analysis**
- Presenting result report and insights using **tableau,**
- Working as a **team leader 3** years during the tenure

04/2019 – 04/2021 Turkey	**Volunteer Experience,** *CoCoDer*

- Workshop event with children on April 23 National Sovereignty Children's Day
- Participation in activities to end child labor.
- Supporting activities to raise awareness of children's rights

EDUCATION

02/2018 – present Turkey	**PhD - Chemical Engineering,** *Ankara University*

Thesis: IoT based control of nano CaCO3 synthesis using machine learning and deep learning algorithms.

08/2010 – 07/2013 Turkey	**Master - Chemical Engineering,** *Ankara University*

Thesis: Accidental risk assessment in atmospheric distillation column in oil refinery

08/2004 – 06/2009 Turkey	**Bachelor's - Chemical Engineering,** *Ankara University*

PROJECTS

Churn Prediction for Bank Customer, *Tensorflow, Keras (ANN)* ⬀
A model that reflects whether a bank's customers leave the bank (close their accounts) or remain customers.

Cancer Classification, *Tensorflow, Keras (ANN)* ⬀

Clothes Classification, *Tensorflow, Keras (CNN)* ⬀

Malaria Data Classification, *Tensorflow, Keras (CNN)* ⬀

Cat-Dog Classification, *Tensorflow, Keras (CNN)* ⬀
The Dogs vs. Cats dataset is a common computer vision dataset in which pictures are classified as either including a dog or a cat.

Cancer Prediction, *K-Nearest Neighbors-Pipeline* ⬀

Heart-Stroke Prediction, *Logistic Regression, KNN Classifier,Yellowbrick* ⬀
With LogisticRegression, I created a model that can predict whether a patient is likely to get a stroke based on features like gender, age, various diseases, and smoking status.

Whether Prediction,
AdaBoosting, Gradientboosting, XGBoostig, Decision Tree, Logistic Regression, SVM, KNN ⧉
Prediction model for whether a particular internet user clicks on an Ad on a company's website.

Slump Flow Modeling, *Linear, Lasso, Ridge, ElasticNet Regression* ⧉
Modeling slump flow of concrete using Advanced Regression Parameters.

AutoScout Car Price Prediction Model, *Linear, Ridge, Lasso Regression* ⧉
An original dataset taken directly from the given web page was used. With Dataset, a Machine Learning model was developed to predict the price of cars based on their given characteristics. You can find more details in the notebook for this project, where I tried to reach the most precise result with the least number of features.

Tree types prediction, *SVC,Decision Tree, XGBoost, Random Forest* ⧉

Customer segmentation, *Random Forest, PCA* ⧉

Car Pricing Model,
Random Forest Regression, AdaBoosting, GradientBoosting, XGBoosting Regressor ⧉

Diabetes Analysis and Modeling, *SVM* ⧉

AutoScout Price Prediction Deployment, *Xgboost, Streamlit,Sklearn, AWS* ⧉

The Relationship Between Gender and Policing Project, *Pandas- Matplotlib- Seaborn* ⧉

Customer Analysis, EDA, *Pandas- Matplotlib-Seaborn* ⧉
Exploratory Data Analysis & Viz. on Customer dataset.

Bike Demand Project, *Pandas, Matplotlib, Seaborn* ⧉

Customer Retention Analysis, *MS Sql Server* ⧉

Nintendo Switch Company Sales Analysis, *Tableau* ⧉

Covid-19 Tracking, *Designed a Tableau Dashboard that focuses on cumulative and new cases/deaths globally for first quarter of 2022. -Tableau* ⧉

Exploratory Data Anlaysis with Autoscout, *Pandas, Matplotlib, Seaborn* ⧉
Car price estimation consists of 3 steps of EDA Project: Data cleaning, Handling Missing Values, Handling outliers

Insurance Charges Prediction Model Deployment, *Linear Regressor(Lasso, Ridge), Random Forest Regressor, AdaBoosting, XGBoosting Regressor; Deployment-Streamlit*

Insurance charges estimation model according to people's age, bmi, smoking, etc.

CERTIFICATES

- Machine Learning with Python - Clarusway ⬈
- Data Visualization with Tableau - Clarusway ⬈
- Statistics - Clarusway ⬈
- Data Analysis with Python - IBM ⬈
- Databases and SQL for Data Science with Python (HONORS) - IBM ⬈
- Data Visualization with Python -Clarusway ⬈
- Deep Learning -Clarusway

PUBLICATIONS

30/12/2020	**Data Sources Used For Risk Analysis in Industrial Facilities and a Sample Case Study,,** *Eskişehir Osmangazi Üniversitesi Mühendislik ve Mimarlık Fakültesi Dergisi.* ⬈
13/04/2014	**Accidental Risk Assessment on Atmospheric Distillation Column in Oil Refinery,,** *Chemical Engineering Transactions* ⬈