HAMZA MURTAZA

Data Scientist & Machine learning Engineer

+923330223551

hamzamurtaza11@gmail.com

www.hamzamurtaza.com

github.com/HamzaMurtaza11

linkedin.com/in/hamza-murtaza

PERSONAL STATEMENT

- Love to play with data, analyzing and visualizing it to explore valuable insights from the data.
- Having a passion to understand, preprocess and train the data with suitable evaluation and tuning.
- Having a problem-solving and innovative mind with an excitement to grow and enhance data science and machine learning skills.

EDUCATION

- **NED UNIVERSITY OF ENGINEERING AND TECHNOLOGY, KARACHI, PAKISTAN.**
  Bachelors Of Engineering In Software Engineering, 2018-2022.
- **ADAMJEE GOVERNMENT SCIENCE COLLEGE, KARACHI, PAKISTAN.**
  Intermediate In Pre-Engineering, 2016-2018.
  Percentage: 86%
- **MAK WAY GRAMMAR SCHOOL, KARACHI, PAKISTAN.**
  Matriculation, Computer Science.
  Percentage: 89%

PROJECTS

HEART DISEASE DETECTION SYSTEM

1. This is Artificial Intelligence based **Binary class classification** project coded using **scikit-learn**. The problem is to find out whether a person or patient has a heart disease or not. This artificial intelligent program (machine learning) can be a fantastic solution to figuring out whether a patient has a heart disease by using complex programming and statistical techniques.
2. The original data set of heart disease prediction containing 14 attributes with 13 clinical attributes and one target attribute is taken from **UCI MACHINE LEARNING REPOSITORY**.
3. Packages and frameworks used in this project are **PANDAS** (python package for data analysis and data frame construction), **NUMPY** (python package for numerical operations), **MATPLOTLIB** and **SEABORN** for plotting and data visualization, **SCIKIT-LEARN** for machine learning, modelling, evaluation, tuning and experimentation.
4. The data is pre-processed, removing all null values, converting non-numeric data into numeric form using one-hot encoding technique and done feature scaling to normalize and standardize the data.
5. Strong **Exploratory Data analysis (EDA)** is performed including Descriptive analytics (mean, standard deviation etc.), correlation matrix between dependent and independent variables to find out relation between attributes and also identify important attributes in the dataset.
6. The classification models such as **Random-Forest**, **K-Neighbors** and **Logistic regression** classification models are coded to compare which model gives the best result for heart disease data training.
7. Hyper-parameter tuning of model is performed through **GridSearchCV** and **RandomizedSearchCV.**
8. The Evaluation metrics (accuracy, ROC/AUC, confusion matrix, f1 score, precision) are measured.

DOG BREED IDENTIFICATION SYSTEM

1. This is an Artificial Intelligence based Deep learning project that builds an end to end multi-class image classification model using **TensorFlow 2.0** and **Tensorflow Hub**.

2. This project is used to predict the particular breed of a dog when an image of dog is provided.
3. The data set contains around 20,000 images of dogs of 127 different dog breeds.
4. The data set for deep learning is taken from Kaggle's Dog breed identification competition.
5. Pre-processing of images is performed containing conversion of images into tensors, normalization of images, resizing of images into (224,224).
6. Creation of batches of images with a batch size of 32.
7. The sequential **Keras** layers are created with input layer is the part of **transfer learning**, and for dense layer **softmax** activation function is used with **Adam** optimizer.
8. **TnesorBoard** (Tensorflow's visualization tool kit) is used that provides the visualization and tooling needed for machine learning experimentation for example tracking and visualizing metrics such as loss and accuracy.

Check my GITHUB profile for more projects.

SKILLS

MACHINE LEARNING/DEEP LEARNING	DATA ANALYTICS
• Scikit-Learn • Tensorflow and Tensorflow Hub	• Numpy and Pandas • Microsoft Excel
DATA VISUALIZATION	WEB-FRAMEWORKS
• Matplotlib • Seaborn • Bokeh	• Flask • Django • HTML/CSS • WordPress
DATA BASE	GRAPHIC DESIGNING
• Microsoft SQL server • Mongo DB • Microsoft Access	• Adobe Photoshop • Adobe Illustrator
UI/UX DESIGNING	SOFT SKILLS
• Figma	• Documentation (research papers, reports) • Presentation • Working with Teams • Time Management • Communication Skills

PROGRAMMING LANGUAGES

- Python (Expert)
- SQL (Expert)
- C/C++ (Basic)

CERTIFICATION

- Microsoft Technology Associate (98-381: MTA: Introduction to Programming using Python).
- Complete Machine Learning & Data Science Bootcamp 2022 (UDEMY).
- Machine Learning at National Centre Of Artificial Intelligence (NCAI), NEDUET.
- Complete Python Developer in 2022 (UDEMY).
- What is Data Science by IBM (COURSERA).
- Graphic Designing from Pakistan Freelancing Training Program.

ACHIEVEMENTS

- Awarded as a City Winner at National Idea Bank Competition.
- Got HEC Scholarship from NED University of Engineering and Technology.
- Published a research paper entitled "IoT in Smart Agriculture: A Systematic Literature Review" at 7th International Zeugma Conference On Scientific Research.