EREN BOZKURT

DATA ANALYTICS

EXPERTISE

Python

- Strong understanding of Python syntax, data types, control structures, and object-oriented programming (OOP) concepts
- Familiarity with Python's built-in libraries for file handling, regular expressions, and working with APIs
- Proficiency in using Python for task automation, including writing scripts to streamline repetitive processes or data manipulation tasks.

Machine Learning

- Strong understanding of foundational statistics and probability theory
- Proficiency in Python libraries (e.g., NumPy, Pandas, Matplotlib, Seaborn)
- Expertise in machine learning algorithms with a solid grasp of their mathematical and statistical foundations (e.g., linear regression, logistic regression, decision trees, support vector machines, random forests, gradient boosting models)
- Familiarity with popular deep learning libraries such as TensorFlow, Keras, and PyTorch, and experience working with neural network models
- Proficiency in evaluating and validating models using metrics such as accuracy, precision, recall, F1 score, ROC curve, and AUC
- Experience in model selection and hyperparameter tuning techniques (e.g., cross-validation, grid search, random search)
- Knowledge and skills in data splitting, creating training-validation-test sets, and evaluating overall model performance.
- Proficiency in interpreting and understanding model results, including feature importance ranking, feature effects, and the decision-making process of the model.

Hadoop

- Experience in file storage and management on Hadoop Distributed File System (HDFS)
- Capabilities to work with big data processing frameworks such as Hadoop MapReduce or Apache Spark
- Basic knowledge and skills about Hadoop Cluster administration and configuration and other components within the Hadoop ecosystem (e.g. Hive, Pig, HBase)
- Using core capabilities for data analysis and processing on Hadoop Streaming

GitHUB

- Project-based working and code management skills
- Experience using version control systems (git) and GitHub
- Ability to use GitHub effectively for code sharing and collaboration
- Ability to contribute or actively participate in open-source projects and to create well-written and understandable documentation (readme files, project documentation

Linux

- Experience working and managing on Linux operating system (e.g. Ubuntu, CentOS)
- Competence in terminal commands and shell scripting
- Knowledge and experience in installing, configuring and updating Linux-based servers and use data analytics tools (Python, SQL etc.) on Linux
- Installation and configuration of commonly used Linux-based tools (e.g. Hadoop, Spark) for data analytics
- Ability to implement and manage security measures (firewall, permissions, secure connections) on Linux and to analyze and debug log files
- Ability to automate data analysis processes in a Linux environment

AWS Cloud Technology

- Experience in file storage and management on services such as AWS S3,database services (Amazon RDS, DynamoDB etc.) on AWS and services such as AWS EMR or AWS Glue
- Knowledge and experience in virtual server management on services such as AWS EC2
- Access to AWS services and basic configuration capabilities using the AWS Management Console
- Using core features for data analysis and processing on AWS (e.g. Amazon Athena, Amazon Redshift Spectrum)
- Capabilities to create user and access roles with AWS IAM (Identity and Access Management) service

SQL&MongoDB

- Strong Database Capabilities
- Data Management and Manipulation
- Data Query and Analysis
- Performance Optimization
- Database Design / Database Integratio