Luis Anton

Senior Data Scientist

Data Scientist with 10 years of experience collaborating with cross-functional teams and ensuring the accuracy and integrity around data and actionable insights. prepared to solve any problems in predictive modeling and insight reporting to boost business efficiency, strategic goals, and profit. Highly experienced in big data and have confident especially in mathematical problems. Possessed strong cooperation and communication.

✉ greywolf03054@gmail.com   📍 Bremen, Germany

WORK EXPERIENCE

Senior Data Scientist
Kalinax

05/2017 - Present

Achievements/Tasks

- Analyzed and visualized dozens of Big Datasets with spark, R, tableau and provided insights to machine learning engineers so that they could develop and manage ML models that made **120K+** revenue
- Built out business datasets to structured and unstructured ones using Hadoop, MongoDB, SQL that provided insights to ML engineers
- Developed Iris based Lung Diagnosis System using CNN deep learning model Collaborating with machine learning engineers and Computer Vision engineers that accuracy is up to 98.5% and saved 10K+ persons so far
- Upgraded and analyzed company's BI database using SQL and tableau that reduced size of company database to 72% and saved annual investment for about 150K
- Partnered with technical team to build out Fake Currency Detection System using KNN model that detected 97.5% fake currency

Data Scientist
StatsBomb

11/2015 - 05/2017

Achievements/Tasks

- Cleaned and integrated large sets of structured and unstructured datasets using pandas and postgreSQL from the ground up for use in ML models and products that reduced size of dataset to 87% and improved model accuracy to 94.5%
- Worked with BI team and stakeholders to analyze and visualize their sport dataset resulting in $480K in annual incremental revenue
- Analyzed and visualized airline data with data analysis team using pyspark, pandas and R improving airport traffic management system

Data Analyst
Idoven

08/2013 - 11/2015

Achievements/Tasks

- Built out marketing data infrastructure from the ground up using SQL to provide real-time insights into product
- Created and prepped data from NET to help data scientists build marketing mix models that lifted ROI by 4 basis points

Data Engineer Internship
pearson and associates

07/2012 - 08/2013

Achievements/Tasks

- Supported building structured and unstructured databases of companies using SQLite
- Analyzed several marketing datasets using R, SQL and tableau that increased 76K annual income

SKILLS

Statistics   Probability   C++   C#

Python   R   Spark   Matlab   SQL

Tableau   SQLite   kafka   redshift

PostgreSQL   SQL Server   Hadoop

Big Data   AWS Lambda   Plotly   NoSQL

Scala   machine learning   MongoDB

Docker   Tensorflow   NLP

AWS SageMaker   AWS EC2   AWS S3

Pandas   Matplotlib   Numpy   PyTorch

GCP   scikit-learn   Matlab   Cloud

Apache Flink   Cassandra   Hive   airflow

Talend   jupyter

PERSONAL PROJECTS

Pneumonia Diagnosis System using Iris recognition (07/2021 - 07/2022)

- Prepared and parsed Iris dataset to improve accuracy 98.5% and provided insights of machine learning model developing

Fake Currency Detection System (07/2020 - 11/2020)

- Using only 10K dataset, achieved over 99.87% accuracy on train dataset and 97.5% accuracy

Resume Parse System (04/2020 - 06/2020)

- Parsed and visualized resume of applicants so that ML engineers select more accurate model

EDUCATION

Bachelor of Mathematics
University of Tsinghua

09/2006 - 04/2010

Bachelor of computer science
University of Tsinghua

09/2010 - 04/2012