BASIL K JOSE

DATA SCIENCE/MACHINE LEARNING ENGINEER

basilkjose123@gmail.com
+91 9746098081
Wayanad
linkedin.com/in/basil
github.com/basilkjose

Over 2.5 years of experience as a Data Science/Machine Learning Engineer with a proven track record of successful projects and a deep passion for leveraging data-driven insights to solve complex problems.

PROFESSIONAL SUMMARY

- Proficient in programming languages such as Python and SQL.
- Strong background in statistical modeling, data analytics, machine learning and deep learning techniques.
- Proficient in data visualization, data cleaning, preprocessing, and feature selection.
- Expertise in statistics analysis, analyzing data, performing data analytics, making data-driven decisions, and utilizing predictive modeling techniques.
- Proficient in computer vision techniques including image classification, segmentation, and object detection with a deep understanding of image processing algorithms.

WORK EXPERIENCE

MACHINE LEARNING ENGINEER

OptiSol Business Solutions

Dec 2021 - Present *Chennai, Tamilnadu, India*

Achievements/Tasks

- Develop and implement computer vision solutions for diverse use-cases, including power pole maintenance and infant diaper-size prediction.
- Prepare and analyze data for machine learning models, ensuring high-quality and reliable inputs.
- Design and develop Backend APIs for training and inference pipelines in object detection and segmentation, utilizing Flask APIs to create efficient micro-services.
- Utilize advanced python file handling techniques to effectively process and prepare data, including working with JSON and XML files.
- Collaborate with cross-functional teams and stakeholders to gather and understand business requirements, ensuring successful implementation of solutions.

DATA SCIENTIST

Apes AI

Jan 2021 - Dec 2021 *Kochi, Kerala, India*

Achievements/Tasks

- Detected fire, smoke, and personal protective equipment using yolov5 models, ensuring accurate and reliable detection results.
- Conducted Exploratory Data Analysis (EDA), performed preprocessing, and developed predictive models to gain insights and drive decision-making.
- Applied feature engineering techniques to create new features, enhancing the performance and effectiveness of predictive models.
- Collected open-source data and filtered data based on problem statements, ensuring the availability of relevant and high-quality data for analysis and modeling.
- Developed a model to detect Not Safe For Work (NSFW) images, blur and blank images, leveraging existing solutions.

SKILLS

Python | SQL | Machine Learning | NLP | Deep Learning | Data analysis | PyTorch | Tensor flow | Computer Vision | Keras | Pandas | NumPy | Scikit-learn | OpenCV | Matplotlib | Seaborn | Plotly | Git | Jupyter Notebook | Visual Studio

EDUCATIONAL BACKGROUND

Cochin University (CUSAT) **2019**
BTECH COMPUTER SCIENCE AND ENGINEERING

Board of Higher Secondary Examinations, Kerala **2014**
HIGHER SECONDARY

Board of Public Examinations, Kerala **2012**
HIGH SCHOOL

CERTIFICATES

- Applied AI course
- Python for Data Science and Machine Learning Bootcamp - Udemy
- SQL For Data Science - Udacity
- Version Control with Git - Coursera
- Introduction to Machine Learning in Production - Coursera

PUBLICATIONS (BLOGS)

- Twint: Twitter Scraping Without Using Twitter's API
- Time Series Forecasting using LSTM
- Natural Language Preprocessing: Steps for Text Data Preprocessing
- Reducing Commercial Aviation Fatalities

PROJECTS

Automated Utility and Power Pole Maintenance
Technology: Python, RESTful APIs (Flask), MySQL, Keras, TensorFlow, PyTorch and OpenCV

- Developed a computer vision platform for image segmentation and detection, including instance, semantic,YOLO and MM Detection.
- Engineered a logic to measure power pole violations using the output of segmentation models.
- Calculate visual distance measurements capabilities like getting the pole height using pixels, finding the distance between one object to another, measuring the top and bottom wire etc.,

Insurance Lead Prediction
Technology: Python, Sckit-learn, Random Forest, XGBoost

- Cross-sell health insurance to the existing who may or may not hold insurance policies with the company.
- A machine learning predictive model needs to predict whether or not a customer will purchase a policy.

Blood Group Demand Forecasting
Technology: Python, Data analysis, Prophet, LSTM

- We want to forecast demands for each blood group for the next one week using historical data. That way, we can prevent both wastage and shortage of blood bags.
- We use both statistical and advanced LSTM models for forecasting.

Personal Character Analysis Using NLP From Twitter Data
Technology: Python, Twint, Gensim, TextBlob

- Scraped user tweets using Twint based on various filters.
- Preprocessed the scraped data for unsupervised learning using gensim.

LANGUAGES

English — *Full Professional Proficiency*

Malayalam — *Native Speaker*

INTEREST

Travelling | Movies | Sports