ASHISH SIMON HARRISON

Indiana - 46815

+1 2602650175

LinkedIn - ashish-simon-harrison-270803115

ashishsimonharrison@outlook.com

PROFESSIONAL PROFILE

As a Data Scientist with 2+ years of industry experience, I specialize in Machine Learning, Deep Learning, and NLP applications. I have a strong background in Python and experience deploying end-to-end ML projects using Django, Flask, and SQL. My expertise includes creating and training models using Neural Networks, Keras and TensorFlow, as well as working with OpenCV and Optical Character Recognition using PyTesseract and AWS Tesseract. In addition, I am proficient in various Python libraries, including NumPy, Pandas, SpaCy, NLTK, PIL, TensorFlow, and SAaS code.

Skills

Data Structures, Algorithms, Data Science, Dataset Wrangling, Machine Learning, Neural Networks, Statistics, Deep Learning, Computer Vision, Python, OCR, Data visualization, Power BI, Django, Flask, SQL, Postgres, Keras, Tensorflow, NumPy, Pandas, OpenCV, Tensorflow, Cuda, PIL, SpaCy, Scikit, Sklearn, NLTK, Java, C, C++,HTML, CSS, JavaScript, Php, Github, KNN, KMeans, Naïve Bayes, Decision Trees, REST APIs,Docker, BERT, BART, Transformer Models.

WORK EXPERIENCE

Capgemini, India
Associate Consultant   **June 2021 - July 2022**
- Conducted exploratory data analysis techniques to identify patterns, relationships, and trends, resulting in actionable insights.
- Validated data for multiple international projects (UK, Middle East, USA), ensuring accuracy and completeness of data sets.
- Engineered a handwritten text recognition model using Keras and TensorFlow(Attention Mechanism), achieving a recognition accuracy.
- Implemented a multi-class classifier using Principal Component Analysis and ResNet to improve classification accuracy.
- Designed session management and encryption API's in Django, ensuring secure communication and data privacy using REST APIs.
- Extracted unstructured data from AutoCad PDF's using AWS Textract, YOLO CV Model, and OpenCV, increasing data accessibility and improving workflow efficiency.

Capgemini, India
Senior Analyst   **July 2019 - May 2021**
- Developed and optimized Deep Learning models using TensorFlow and Keras.
- Applied advanced algorithms and programming skills to create accurate predictive models.
- Expertise in Natural Language Processing (NLP) using transformer models and spaCy.
- Utilized OpenCV for noise removal and Image Processing techniques.
- Extracted data from structured and unstructured files using OCR tools such as Pytesseract.
- Developed end-to-end applications in Python using containerization with Docker and web frameworks like Django.
- Worked on multiple transformer models like BERT and BART and used techniques to enhance the pretrained models.

PROJECTS

Form Interpreter and Analyzer for Insurance Claims
- Optimized Handwritten Text recognition model to improve accuracy.
- Created multiple APIs for image processing, data communication, and post-validation.
- Spearheaded the development of a multi-class classifier using Principal Component Analysis and ResNet.
- Developed the Knowledge Base for post-validation to improve accuracy and efficiency.
- Designed and implemented session management and encryption API for enhanced data security.

Risk Regulatory Dashboard
- Utilized Natural Language Processing (NLP,NLTK) techniques to analyze documents and websites and determine their sentiment.
- Extracted data from PDFs and web pages using the BeautifulSoup library.
- Conducted thorough data cleaning procedures and generated a word cloud showcasing key terms.
- Developed a comprehensive statistical report detailing the sentiment of the analyzed documents and deployed it using docker.

SAS to Python
- Managed and led a team of 15+ members.
- Created an algorithm that breaks down SAS code into multiple blocks.
- Developed a mechanism to convert SAS blocks to Python blocks using Object Oriented design and Python.
- Designed and implemented an auto-integrator to seamlessly integrate different code blocks.

DocumentAI
- Utilized OpenCV to preprocess images.
- Applied a neural network classifier using NasNet and ResNet to classify images.
- Extracted required information from documents by drawing contours using OpenCv and using PyTesseract.

EDUCATION

Purdue University, Indiana, USA
Master of Science in Computer Science   **August 2022 - May 2024**
                                                                                                                                                GPA – 3.83

SRM Institute of Science and Technology – Chennai, India
Bachelor of Science Computer Science and Engineering   **July 2015 - November 2019**
                                                                                                                                                GPA - 8.92
- Presented a paper on IoT (Internet of Things) Home Automation Technology using Audio Commands (Major Project)
- Local Navigator - Real time Navigation App for college premises using python (Minor Project)
- SentiSouls : Photography and poetry blog (Extra curricular)

CERTIFICATIONS

- Second runner up in Global Data Science Challenge Hackathon 2021
- Machine Learning certification by IBM (Coursera)
- Machine Learning Deep Learning OpenCv