Akshay Sharma

Data Science Developer

Data Science enthusiast with 4 years (approximately) of work experience in various Data Science domains such as Artificial Intelligence, Machine Learning, Deep Learning and NLP.

Contact

Address

1515, Preet Vihar, Roorkee, Uttarakhand, 247667

Phone

+91 6396674227

E-mail

akshaykr.sharma@outlook.com

LinkedIn

Akshay Kumar Sharma

Skills

Deep Learning

Very Good

Machine Learning

Very Good

Python

Very Good

Power BI

Good

Natural Language Processing (NLP)

Very Good

Named Entity Recognition (NER)

Good

Certifications

- TensorFlow Developer: Professional Certificate - Coursera (Certificate)
- Python For Data Science: Professional Certificate - IBM (Certificate)
- Applied Machine Learning - Applied AI (Scaler Academy)
- Data Science - Internshala ()
- Docker Hands On - ChrysalisGold
- Introduction to Linux - edX

Blogs

- Steps to deal with Overfitting and Underfitting on Image Data using Image Augmentation. (Blog URL)
- Maps with Python Using Folium (Blog URL)
- Introduction to Neural Networks and Activation Functions (Blog URL)

Work History

Data Engineer (Data Science)
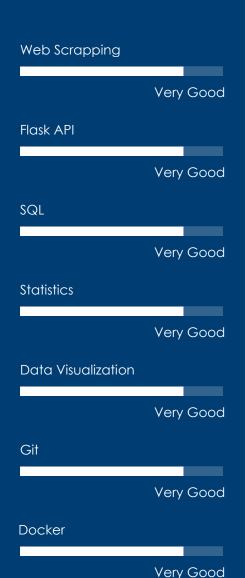
ChrysalisGold, Bengaluru, Karnataka

05/2021 – Present

- **Technologies and Skills Used:** OCR, NLP, Machine Learning, AWS, Docker, Python, Flair, NER, Data Preprocessing, Data Postprocessing, Azure, Client Presentations, Data Engineering, Research and Development, Git.

AI Developer

Vaibhav Global Ltd, Jaipur, Rajasthan

10/2020 – 04/2021

- **Technologies and Skills Used:** Web scraping, Power BI, PostgreSQL, NLP, Deep Learning, Model Building.
- Power BI Reports to give a higher understanding of the given data to the business teams.
- Web Scraping using Selenium and Beautiful Soup.
- PostgreSQL is used extensively while working with inserting or fetching data from the database.
- Deep Learning used to build some small projects and to enhance the existing projects like image segmentation.
- Machine Learning used to build a recommendation system from scratch.

Skills

Web Scrapping

Very Good

Flask API

Very Good

SQL

Very Good

Statistics

Very Good

Data Visualization

Very Good

Git

Very Good

Docker

Very Good

Data Scientist

Prudle Labs Pvt Ltd, Bangalore, Karnataka

04/2019 – 09/2020

- **Technologies and Skills Used:** Product Development using NLP, Machine Learning, Deep Learning, Model Deployment and Data Visualization GCP, AWS, SQL. Mongo DB and POC's Development.
- Product Development using NLP, Machine Learning, Deep Learning, Model Deployment and Data Visualization.
- POC's for various Deep Learning and Machine Learning projects.
- Solved Translation Memory utilization problem using NLP approach.
- sentiment analysis model using Machine Learning.
- OCR to extract texts from scanned pdfs.
- Transfer Learning from Mozilla DeepSpeech model to convert audio files into text format.
- Flask APIs to deploy models and also give them a UI. GCP and AWS in some instances where high end systems needed.

Projects

- **AUDIO TO TEXT TO SENTIMENT ANALYSIS:**
  I am the head of project in this implemented from scratch to 100% deployment right from research and development to model selection to deployment using flask API.

- **Automatic Transcription of videos**:
  I am the project head developed till PoC presented to Zee networks this project takes input any video and create subtitle file for that using Deep Learning and audio noise cleaning.

- **OCR RPA**:
  I am part of team (3 members) who created this solution of OCR that takes tiff files as input and extract required data from those files these extracted information is stored in csv's which than go through complex post processing where we ensure final csv to meet all client's requirements delivered more than 2 Lakhs files output to client till now.

- **Utilizing Translation Memory (NLP)**:
  I am the project head. This was an internal project in which we utilized translation memory (database of translation done in past) to predict the upcoming translation using NLP.

- **Grantor Grantee Extraction (NER)**:
  I was the project head responsible for research and development to extract grantor and grantee from tiff files document using deep learning. After extensive research I implemented Flair Model (NER Model) to extract the required fields and populate the csv for the same we got 85% accuracy with this.

- **ASPEN DEPLOYEMNT**:
  This is the company's hot product. I was responsible for deployment of all the models involved using Docker, Kubernetes on AWS. I had successfully

deployed this product on various servers and carried out client presentations for the same. I had also deployed this product on AWS (Kubernetes).

- **Database Adapter**:
  I developed a python-based database adapter that can fetch and store data from various databases like MySQL, Oracle and many more. This is highly needed to get the data so that can be used to predict and train models.

- **Tabular Data Extraction**:
  This is R&D project in which we tried to extract the tabular information from tiff files which have complex tabular structure. We Researched with various things like AWS Textract, RetinaNet, Camelot. We did the PoC using Camelot and RetinaNet.

Education

Bachelor of Technology: Computer Science and Engineering

College of Engineering Roorkee - Roorkee, Uttarakhand | 73.33%

08/2014 - 05/2018