

## Arvind Kumar

Location: Shahjahanpur, Uttar Pradesh, 242001

Email: stdntlfe@gmail.com

Mobile: +91-9792979142

LinkedIn: <https://www.linkedin.com/in/stdntlfe>

GitHub: <https://github.com/stdntlfe>

## Professional Summary

Experienced Data Scientist with a strong foundation in Machine Learning, Deep Learning, and AI solutions. Skilled at building and deploying large-scale data-driven applications using Python, TensorFlow, FastAI, and other cutting-edge tools. My expertise spans across computer vision, NLP, image processing, and predictive analytics, with a proven track record in delivering impactful solutions for food, healthcare, and telecom industries. Proficient in leveraging frameworks like Flask and FastAPI, and big data platforms for crafting effective business intelligence tools and data-driven insights.

## Technical Skills

- Programming Languages: Python, R, SQL, Bash
- Frameworks: TensorFlow, FastAI, MXNet, Keras, Scikit-learn, PyTorch
- Web Development: Flask, FastAPI, HTML/CSS, JavaScript
- Databases: MySQL
- Cloud: AWS, GCP
- Data Visualization: Tableau, Power BI, Matplotlib, Seaborn
- Machine Learning: SVM, Decision Trees, Random Forest, XGBoost, k-NN, K-Means
- Deep Learning: CNNs, RNNs, GANs, Transfer Learning, ImageNet
- NLP: SpaCy, NLTK, Transformer Models (BERT, GPT), Tesseract OCR
- Other Tools: Docker, Kubernetes, Git, Jenkins

## Professional Experience

### Senior Data Scientist

#### Meraki Training Solutions (Contract) | October 2020 - October 2024

- Spearheaded the design and implementation of machine learning models for complex image processing tasks, achieving significant accuracy improvements and error reduction.
- Led the development of a personalized recommendation system that boosted customer retention by 20% in the food sector.
- Integrated AI-driven predictive models with business intelligence tools, optimizing operational workflows.
- Created dynamic web applications for image detection and recognition tasks, serving

diverse needs from vehicle detection to facial recognition.

- Streamlined data pipelines and predictive modeling using Hadoop and Spark, accelerating processes for healthcare and telecom sectors.

### Data Science Engineer

**Dew Solutions Pvt Ltd | Gurugram, Haryana | July 2019 - August 2020**

- Built machine learning and deep learning models for NLP applications, automating customer interactions via chatbot systems.
- Designed a recommendation system to enhance user engagement by 15%, leveraging collaborative filtering.
- Developed large-scale web scraping solutions for healthcare market analysis.
- Implemented machine learning models for sentiment analysis and customer feedback categorization, enhancing feedback analysis.

### Data Science Trainee

**Azure Power | Aerocity, New Delhi | March 2019 - July 2019**

- Analyzed solar energy data for performance optimization of rooftop installations.
- Used time-series forecasting to predict energy generation, supporting proactive maintenance.
- Created dashboards to visualize and monitor key performance metrics for solar installations.

## Projects

### Vehicle Detection and Classification

Developed a scalable, real-time vehicle detection and classification system to support smart traffic management. Leveraging YOLOv7 and Flask, this system can identify vehicle types from live video streams with high accuracy.

**Key Technologies:** Python, YOLOv7, OpenCV, TensorFlow, Flask

- Built a deep learning pipeline, adapting pre-trained YOLO models for specific vehicle categories.
- Optimized model for low-latency processing on edge devices, suitable for real-time deployments.
- Created a user-friendly Flask interface, enabling traffic monitoring and reporting for city authorities.

### Twitter Sentiment Analysis

Designed a sentiment analysis tool for a telecom client, enabling them to monitor customer sentiment on Twitter in real time. This solution provided actionable insights for brand and customer engagement strategies.

**Key Technologies:** Python, NLTK, SpaCy, Scikit-learn, Flask, Twitter API

- Developed a pipeline to collect and preprocess tweets using the Twitter API.
- Implemented sentiment classification using machine learning algorithms and deployed the model in a dashboard.
- Delivered sentiment analysis reports, helping the client make data-driven improvements in customer interaction.

### Face Recognition System

Built a robust face recognition system for a healthcare facility to enhance security by authorizing personnel access to secure areas.

**Key Technologies:** Python, CNN, OpenCV, Keras, TensorFlow, Flask

- Developed a CNN for face recognition, optimized for accuracy under various lighting conditions.
- Integrated with OpenCV and Flask for real-time capture and recognition.
- Provided security features like multi factor authentication and face embedding storage for data protection.

### Speech-to-Text Model Development

Created a multilingual speech-to-text model to transcribe audio for customer service applications, enabling faster responses.

**Key Technologies:** Python, Hugging Face Transformers, PyTorch, Wav2Vec2

- Fine-tuned Wav2Vec2 for accurate transcription across multiple languages.
- Integrated noise-reduction techniques and deployed via a Flask API, streamlining audio-to-text conversion for service applications.

### LLM Integration for Text Generation

Integrated large language models to automate text-based customer interactions, reducing manual intervention and improving response quality.

**Key Technologies:** Python, Hugging Face Transformers, PyTorch, GPT-2, BERT

- Fine-tuned GPT-2 for specific telecom-related queries.
- Integrated the model into an existing chatbot system, boosting efficiency by 50% and improving customer satisfaction.

### Web Scraping for Justdial

Developed a web scraping solution to extract business data from Justdial, providing insights for market analysis and competitive trends.

**Key Technologies:** Python, BeautifulSoup, Scrapy, Pandas, Flask

- Implemented a flexible pipeline for scraping business data with custom filters.
- Enabled structured data export for analysis, supporting data-driven decisions for clients.

### Generative AI for Text-to-Image

Created a text-to-image generation tool using Stable Diffusion, helping clients automate visual content creation for marketing.

**Key Technologies:** Python, Stable Diffusion, Hugging Face Diffusers, PyTorch

- Fine-tuned Stable Diffusion for high-quality image generation.
- Built a web interface in Flask, allowing users to input text and receive custom-generated images in real time.

### Education

Bachelor of Technology in Computer Science

Lovely Professional University | Jalandhar, Punjab | 2019

### Certifications

Python Data Scientist Certification (DataCamp) – 2018

Deep Learning Specialization (Coursera) – 2017