

# Pawan Kalyan Jada

Mobile: 9390199603

LinkedIn: [linkedin.com/in/pawan-kalyan-9704991aa/](https://www.linkedin.com/in/pawan-kalyan-9704991aa/)

Email: [pawankalyanjada@gmail.com](mailto:pawankalyanjada@gmail.com)

Github: [github.com/JadaPawanKalyan](https://github.com/JadaPawanKalyan)

## EDUCATION

- Indian Institute of Information Technology Tiruchirappalli** Trichy, India  
*Bachelor of Technology - Computer Science; CGPA: 8.7/10* 2019 - 2023

## EXPERIENCE

- Celebal Technologies** Remote  
*Data Science Intern* Jan 2022 - Jun 2022
  - Tabular Question Answering:** Improved the performance and reduced the latency of Tabular Question Answering application by using state of the art ML approaches.
  - Product Recommendation:** Optimized the performance of a Product Recommendation engine by using BERT and cosine similarity leading to over 25% increase in accuracy of predictions.
  - Call Summarization Engine:** Devised and Implemented an architecture for a call summarizing system, that generates summary of the call based on transcripts and various other KPI's measured during the call conversation.
- Classforma AI** Remote  
*Data Science Intern* Sep 2021 - Dec 2021
  - Document Retrieval:** Assisted with the development of a search engine system for retrieving relevant documents from a collection of documents using key phrases and terms.
  - Automatic Answer Extraction:** Built and improved an extractive Question-Answering model by training it with a newly approached strategy that led to a 7% overall increase in accuracy.
- Drishite™** Remote  
*Machine Learning Intern* June 2021 - Sep 2021
  - Invoice Detection:** Developed an Automatic Invoice Detection System that automatically detects the invoices from a pdf/image, process and stores them.
  - Number Plate Identification Pipeline:** Designed and created a pipeline for detection and identification of number plates from incoming video stream.

## PROJECTS

- Youtube Video Summarizer:**
  - A web app that summarizes what's happening in a video when given a YouTube link. It uses the YouTube Transcript API to get the transcript for that specific video, and then runs the machine learning logic on top of it to generate the summary.
  - The idea is to summarize the most important points of a YouTube video in the shortest time period possible. [Link](#).
- Analysis of Delivery Partners:**
  - It involves doing exploratory data analysis, data preprocessing, visualization, and feature engineering on a delivery partner's dataset with over 130K rows, which includes customer orders, customer zones, order placement time, and many other features.
  - The final model built has 96% Accuracy. [Github](#).
- End to End QA System:**
  - A complete data science project which involves data preprocessing and visualization, model training, evaluating by picking the best checkpoint of the model, creating a web application for the UI and deploying it on Streamlit.
  - The objective is to eliminate the need to search for answers in a lengthy pdf or document, which would require a significant amount of time and human effort. [Link](#).

## SKILLS

- Languages:** Python, C, Java
- Big Data:** PySpark
- Cloud:** MLOps, Microsoft Azure, AWS
- Tools:** GIT, PostgreSQL, MySQL, Docker
- Frameworks:** TensorFlow, Pandas, Flask, Scikit-learn, Matplotlib, SpaCy, OpenCV, Pytorch
- Data Science:** Probability and Statistics, Data Analysis, Predictive Modeling, Model Deployment
- Natural Language Processing:** Word Vectorization, Bag of Words, Cosine Similarity, Classification Models, Naïve Bayes Model, Chatbot, Machine Translation, Document Search.

## ACHIVEMENTS & ADDITIONAL INFORMATION

- Hackathon:** Secured 1st position in the Atrang Hacks'22 hackathon, where we've built a Student helper app which contains components like Text-to-Code, Story Generation etc. [Devpost link](#).
- Problem Solving:** Solved 400+ quality problems on Leetcode([Pawan\\_Kalyan](#)), GeeksforGeeks ([pawankalyanjada](#)) and Hackerrank([pawankalyanjada](#))

## CERTIFICATIONS

- [Azure Data Science Associate](#) - Jan 2022
- [Python Data Structures](#) - June 2020
- [Hackerrank SQL Certificate](#) - Mar 2022