

# Girja Kumari

ML ENGINEER / ML DEVELOPER / DATA SCIENTIST

## Profile

Experienced Machine Learning Engineer who is enthusiastic and committed. Possesses expertise in tasks such as data analysis, ml models designing, training, testing, deployment, and maintaining infrastructure. Dedicated to functioning effectively within a cooperative and optimistic team. Seeking a dynamic environment that offers both challenges to further skill development and the chance to make substantial contributions to the organization.

## Professional Experience

**Experience** – Almost 3 years

### ML Engineer, Senior Member Technical, ADP, Hyderabad

NOVEMBER 2020 — AUGUST 2023

- Employed Transformers, Layout LM, and Large Language Models (LLMs), along with deep learning architectures such as CNN, RNN, LSTM, and attention mechanisms, to extract vital fields/entities that significantly fulfil business requirements.
- Enhanced the effectiveness of the pre-existing BERT framework and flair framework through the implementation of efficient preprocessing and cleansing methods. Created an automated pipeline for retraining models within both frameworks, eliminating the necessity for manual involvement.
- Collaborated with QA and managers to establish and fine-tune objectives. Devised machine learning algorithms to analyse extensive sets of historical data.
- Achieved a reduction of approximately 30% in the indexing expenses for business documents by incorporating the confidence scores from the model's predictions, in addition to the document-level confidence scores from Tesseract.
- Increased auto response for documents from ~36% to ~60% as a result in improvement of performance of models.
- Addressed the intricate issue of model drift by incorporating augmented data into the training dataset, leading to its resolution.

## Education

### M.Tech, NIT Rourkela, Odisha

JULY 2018 — MAY 2020

Post Graduated in Computer Science with 7.8 CGPA.

### B.Tech, Govind Ballabh Pant Institute of Engineering & Technology, Pauri, Uttarakhand

AUGUST 2012 — MAY 2016

Graduated in Computer Science & Technology with 76.66%.

## Personal Details

ADDRESS

A203, Abhee Lakeview, Serenity Layout, Kaikondrahalli, Sarjapur Rd, Bengaluru, 560035

PHONE

8885187014, 8810504421

EMAIL

[singh.grj.01@gmail.com](mailto:singh.grj.01@gmail.com)

DATE OF BIRTH

01/09/1993

## Links

- [IEEE Publication: Effect of Reduced Dimensionality on Deep learning for Human Activity Recognition](#)
- [LinkedIn Profile](#)

## Skills

- Machine Learning, Deep Learning, NLP
- Libraries (TensorFlow, Keras, spaCy, PyTorch, Scikit-learn, Pandas, Numpy)
- Predictive Modelling
- Data Visualization
- ML Algorithms
- Data Structures and Algorithm
- Amazon AWS
- MySQL
- Python, C & C++
- Agile Methodology
- Docker & Kubernetes

## Hobbies

Sketching, Travelling

## **Intermediate, K V NO.1 BEG & C, Roorkee, Uttarakhand**

APRIL 2010 — MAY 2011

Intermediate in Science & Maths with 82%.

## **Matriculation, K V NO.1 Garhi, Udhampur, J & K**

APRIL 2008 — MAY 2009

Passed with 80.6%.

## **IEEE Publication**

### **Effect of Reduced Dimensionality on Deep learning for Human Activity Recognition**

SEPTEMBER 2019 — JUNE 2020, ROORKELA

Published in 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)

Inertial sensors were attached to individuals to record angular velocity and acceleration during outdoor daily and sports activities. A deep learning classifier, a stacked LSTM network, was trained on this data to extract complex features from activity signal sequences.

## **Project**

### **Human Activity Recognition**

- A time series classification problem which predicts movement of a person based on sensor data.
- Collected new dataset of 9 outdoor daily and sport activities using Sparkfun 9 DoF IMU sensor by the involvement of 10 human subjects.
- Used heat map and feature importance feature selection techniques, to identify highly contributing features towards classification.
- Performed comparative study to analyse the effect of reduced length feature vector on the performance of deep learning and machine learning techniques used for HAR based on inertial sensor data.

### **Movie Recommender System**

- Implemented tendency based collaborative filtering for recommending movies and calculated diversity among recommended movies using popularity of movies by using MovieLens Dataset.

### **Content Based Image Retrieval**

- A MATLAB based software application, with an image database, that utilized colour, texture, and shape features of the images in the database as the basis of comparison and retrieval.

## **Extra-curricular activities**

Coordinator of Sport and Cultural events at school and college level.