Binay Kumar **Pradhan**
Data Science Engineer

Bangalore, India - 560076

+91-9556226118

pradhan.binay@outook.com

www.linkedin.com/in/i-binay

https://github.com/VinayIN

https://wezn.github.io/

Collaborative, dependable, and accomplished Data Science Engineer with experience in developing end-to-end data science product. Skilled in research, data analytics and providing solutions to NLP problems using deep learning techniques. Motivated to learn, apply, and adapt to changes generating maximum throughput.

Experience

Data Science Engineer	Utopia India Pvt Ltd - Bangalore, Karnataka	July 2019 - Present

- **4C – Text cleaning and structuring Package:**
  - o Responsible for creating data apps for the clients to use the 4C package on their dataset and providing quarterly KT to the team on the changes made to it
  - o Identified bottleneck and refactored it to a package, integrated it with multiple UI's and visualization to demonstrate capabilities of the tools to client and created API modules to expose each tool of the package
- **Text parsing and information extraction tool:**
  - o Conceptualized reinforcement learning to implement feedback received from team interacting with the tool and demonstrated a POC on the use case of Transformers and BERT model for standard text generation of industrial equipment.
  - o Developed a regex-based parser on top of ML parser along with a consensus algorithm to accurately extract information from text descriptions
- **Named entity mapping tool**: Developed an embedding based entity mapping tool with 85% filling accuracy and 15x better performing than excel macros and rule-based tool used. Also generated few plugins like spell correction, similarity keyword extractor along with the tool to accelerate cleansing activities for developers
- **Database:** Maintenance and manipulation of data to the DB on-demand
- **BOM:** Integrated business rules to the contents of tables extracted from BOM pdf
- Assessed and setup CI/CD tool for the team
- Assisted on the configuration of AWS SPOT instances for multiple projects and load distribution on the development server
- Created scripts to automate data migration, map to company's taxonomy adhering to SAP guidelines and provided induction to interns on the 4C tools

Machine Learning Intern	Utopia India Pvt Ltd - Bangalore, Karnataka	January 2019 – July 2019

- Developed a text matching algorithm using combination of pretrained vectors and custom embeddings along with set relation techniques
- Developed a new data driven algorithm for text structuring and standardization
- Orchestrated the 4C tool using Makefile for batch file processing

Research Intern	Institute of Development and Research in Banking Technology - Hyderabad, Telangana	May 2018 – July 2018

- **Dataset:** Prepared the Hindi mixed-script dataset for training
- **Publication:** Assisted a fellow scholar on an NLP problem of part-of-speech recognition for mixed-script social media texts and documenting the journal "Improving Code-mixed POS Tagging Using Code-mixed Embeddings"

Skills

Deep Learning • NLP • Tensorflow • Keras • Pytorch • Scikit-Learn • Python • Unix • MySQL • Elasticsearch • Streamlit • Django • flask • Jenkins • Library packaging • Product management • Computer Vision • Golang

Education

Bachelor of Technology in Information Technology	International Institute of Information Technology – Bhubaneswar, Odisha	August 2015 – June 2019

Interests

Hackathon • Basketball • Hiking • Language Learning • Sketching