Jillur Rahman Saurav

Software Engineer

☎ (880) 1726536943 | ✉ sauravsust71@gmail.com | in jillur-rahman-saurav

Skills

Core CS	Machine Learning, Data Structure, OOP
Languages	Python, Java
Tools & Technologies	Keras, Tensorflow, Scikit-learn, Docker, Elasticsearch, MongoDB, Mysql, Git

Experience

Pipilika   *Sylhet, Bangladesh*
Software Engineer   *November 2017 - December 2020*

- Researched, designed and implemented scalable Knowledge graph for Bangladesh Government's Portals using K-Means clustering, Nearest-Neighbours and Elasticsearch.
- Worked on Pipilika's Stemmer, Spell-checker and Document Categorizer whiches are freely available at Pipilika's Developers API for research purpose.
- Developed a new generic parser for parsing web pages. For Bangla language it can parse contents, dates, titles, top images from HTMLs and extract summaries with a high accuracy.
- Developed various backends using Java Spring Boot, Flask, Django, Apache Solr, Elasticsearch, Apache Kafka, Redis based on requirements.

Research

- **Newspaper bias analysis**: The main focus of this research was to find different biases [political bias, gender bias, etc.] presented in the Bangladeshi newspapers. Moreover, this research compared the findings with other countries' newspapers.
- **Fundamental NLP problems of Bangla language**: Explored Deep learning and classical machine learning algorithms for solving fundamental problems of NLP in Bangla language, namely **Document classification, Document clustering, Sentiment analysis, Sentence autocompletion, Parts-of-Speech tagging, Named-Entity recognition.**
- **Speech recognition system for Bangla language**: Built a corpus and speech recognition system for Bangla language.
- **Dengue outbreak prediction**: Analyzed newspaper data to find correlations with DGHS data and to predict Dengue outbreaks for Bangladesh.
- **Autonomous vehicle**: Used Computer Vision [tensorflow API] to detect other vehicles, pedestrians, objects, etc. in real time.
- **Covid-19 data analysis**: Performed various statistical analysis on the data obtained by a self-screening tool.

Publications

Word Completion and Sequence Prediction in Bangla Language Using Trie and a Hybrid Approach of Sequential LSTM and N-gram
2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT)   *28-29 Nov. 2020*

Query Expansion for Bangla Search Engine Pipilika
2020 IEEE Region 10 Symposium (TENSYMP)   *5-7 June 2020*

End to End Parts of Speech Tagging and Named Entity Recognition in Bangla Language
2019 International Conference on Bangla Speech and Language Processing (ICBSLP)   *27-28 September 2019*

Bangla Speech Recognition for Voice Search
2018 International Conference on Bangla Speech and Language Processing (ICBSLP)   *21-22 September 2018*

Education

Shahjalal University of Science and Technology   *Sylhet, Bangladesh*
Bachelor of Science in Computer Science & Engineering   *2018*