Arushi Sharma

☎ (+91) 9205858274 | ✉ arushi.net@outlook.com | ⊙ arushi-08 | in arushi--sharma

Education

Delhi Technological University (formerly Delhi College of Engineering) *New Delhi, India*
BACHELOR OF TECHNOLOGY · ELECTRONICS AND COMMUNICATION ENGINEERING · 7.91/10 CGPA *Aug. 2016 - June 2020*

Indian Language School *Lagos, Nigeria*
CENTRAL BOARD OF SECONDARY EDUCATION · 92.5% IN CLASS 12 · 10 CGPA IN CLASS 10 *Mar. 2016*

Professional Experience

Optum Global Advantage · UnitedHealth Group *Hyderabad, India*
ASSOCIATE DATA SCIENTIST · BACKEND PYTHON PROGRAMMER · AI & ML TEAM *Aug. 2020 - Present*
- Developed and maintained Document Processing Pipelines for a Medical Document Indexing Site in Python using Celery, Flask and Pandas.
- Built a Text Similarity feature to find the similar pages between two documents using Cosine and Levenshtein Similarity metrics.
- Built a Date Search feature using regex to extract various date format types from documents.
- Designed Flask APIs to analyze daily Data Inventory Metrics, Document Processing Time, and Disk Storage.

IHS Markit Ltd. *Gurugram, India*
DATA TRANSFORMATION INTERN · ENGINEERING AND PRODUCT DESIGN TEAM *Jun. 2019 - Aug. 2019*
- Developed a Web Crawling Automation tool to navigate through websites and extract information, using Selenium Frameworks of Python. This decreased data extraction time by 22%.
- Programmed data files to a set format, using NumPy operations on Pandas DataFrames, and built a GUI using Tkinter Python package.
- Deployed it as a full-fledged desktop application, using cx_Freeze.

Publications

Ceasing hate with MoH: Hate Speech Detection in Hindi-English Code-Switched Language
ACCEPTED IN INFORMATION PROCESSING AND MANAGEMENT JOURNAL (IMPACT FACTOR: 6.222) *2021*
- Classified social media tweets as not offensive, abusive and hate inducing.
- Extracted features from tweets: TF-IDF Word level vectors, Word Level N-gram Vectors, and employed Machine Learning Classifiers.
- Developed a Language Transliteration Mapper and finetuned mBERT and MuRIL classifier models.
- Compared the empirical results and reported scores higher than state-of-the-art precision, recall and F1 scores.

Feature Enhanced Capsule Networks for Robust Automatic Essay Scoring · Link
PUBLISHED IN ECML PKDD 2021 (A RANKED CONFERENCE): MACHINE LEARNING AND KNOWLEDGE
DISCOVERY IN DATABASES. APPLIED DATA SCIENCE TRACK *2021*
- Scored student essay responses on various questions.
- Proposed a novel architecture with two key components: BERT enhanced CapsNet model and the Feature Extraction model.
- Boosts the overall average Quadratic Weighted Kappa metric to 81%.
- The proposed work successfully scores most of the adversarial text lower.

English-Hindi Neural Machine Translation-LSTM Seq2Seq and ConvS2S · Link
PUBLISHED IN ICCSP'20 IEEE CONFERENCE *2020*
- Cleaned and tokenized Hind-English Corpus using NLTK library and regex.
- Developed the language translator in Python using ConvSeq2Seq and LSTM Seq2Seq deep learning models.
- Compared these models on the basis of BLEU metric score, training time and hardware utilization.

Projects

ICU Requirement Prediction
- Analyzed dataset to find the pertinent patients' information for classifying them into ICU Required or not. *2021*
- Reported best accuracy scores with Logistic Regression model.
- Built Flask APIs rendering HTML templates and dockerised the application.

Sentimental Analysis on Customer Reviews

- Classified dataset into positive and negative reviews.   *2019*
- Performed text preprocessing using NLTK and regex and employed classical ML, CNN and LSTM models.
- Evaluated performance on Accuracy, Precision, Recall and F1 metrics.

Honors and Awards

2021	**Make IT Happen Team Winner for Platform Architecture & Cloud Hosting**, for advancing Optum technology. Selected among 804 nominations	*Optum, India*
2021	**Conducted Introductory Machine Learning Training Session**, for 100+ Office Employees	*Optum, India*
2021	**3 Employee Recognitions for Performance & Compassion**, for professional work in H1 of 2021	*Optum, India*
2019	**Winner for Best Usage of Personalized Wellness Tracking for Health & Fitness**, Smart India Hackathon. Selected among 1000+ applications	*NIT Trichy, India*
2016	**Sportsperson of the Year, Indian Language School**, For winning 40+ sports certificates from Class 9.	*Lagos, Nigeria*

Certifications

2020	**The Data Scientist's Toolbox**, Coursera
2020	**Improving Deep Neural Networks: Hyperparameter Tuning, Regularization and Optimization**, Coursera
2018	**Big Data Analytics with Artificial Neural Networks**, National University of Singapore
2018	**Hands On Training in Big Data and HADOOP**, Hewlett Packard Enterprise
2018	**Machine Learning: Foundations, Regression and Classification**, Coursera

Leadership Experience

TEDxDTU 2019   *DTU*
CORPORATE HEAD OF BUSINESS DEVELOPMENT   *2017-2019*

- Organized one of the best TEDxDTU events, with a line-up of 10 brilliant speakers including Monica Dogra, Karan Wahi, Laxmi Nair.
- Led the Corporate team, and secured sponsorships for the annual event.
- Trusted to be one of the sincere and dedicated members of the Organizing Committee.

VIBE Dance Society   *DTU*
CORE MEMBER   *2016-2020*

- Won prizes in Inter-College dance team competitions: BITS Pilani 1st Prize, IIT Mandi 2nd Prize.
- Organized and oversaw the budgeting of the Annual Streets Dance Competition, Switch the Funk up.
- This event is one of India's best, judged by World Famous Dance Artists, and attended by 500+ dancers from all over India.
- Organized Dance Workshops at University Campus, attended by 50+ new dancers.

School Sports Events   *Nigeria*
SPORTS HEAD   *2015-2016*

- Runner-up in Inter-State Table Tennis Championship representing Lagos State in Nigeria.
- Organized an independent Inter-School Sports Event, heading a committee of 15 students.
- Captained the Girls Basketball School Team at intra and inter sports events.

Skills

Programming	Python · Java · Bash · Git **Frameworks**: Pandas · NumPy · Flask · Celery
ML & NLP	Scikit-Learn · Pytorch · NLTK · Keras
DevOps	OpenShift · Jenkins · Docker **Database:** MongoDB · Redis **Testing:** Selenium