

Jay D. Vala

MASTER OF SCIENCE · DIGITAL ENGINEERING

Weißenburg Str 41a, Würzburg 97082, Germany

☎ (+49) 1520 988 1070 | ✉ jay.vala@protonmail.com | 🌐 www.jayvala.in | 📷 jdvala | 📺 jayvala

"Success comes to those who change with the changing times."

Education

Otto-von-Gurieke University

Magdeburg, Germany

M.Sc. DIGITAL ENGINEERING

October 2015 - July 2019

- Master Thesis titled "Classification of multilingual legal text using deep learning: Evaluation of general purpose resources for legal domain specific task"
- Did university project on Implementing Network Intrusion Detection using Tensorflow and Microsoft Cognitive Toolkit.
- Successfully completed data analytic project to implement several classification and machine learning algorithms using Python.
- Gained a good understanding of various data analytically approaches and other statistical measures.
- Obtained hands-on experience of various Natural Language Processing Tools.
- Course undertook: Swarm Intelligence, Organic Computing, Information Retrieval

GTU(Gujarat Technological University)

Ahmedabad, India

B.E. IN ELECTRONICS AND COMMUNICATIONS

August 2010 - August 2014

- Courses undertook: Digital Communication, VLSI, Basic and Advance Electronics
- Final Year Project: Quad-Copter Surveillance System.

Experience

Scoutbee GmbH

Würzburg, Germany

DATA SCIENTIST

April 2021 - Present

- Together with the team, trained and deployed models for Entity resolution.
- Benchmark different models using MLFlow
- Manage different version of data using DVC.

Scoutbee GmbH

Würzburg, Germany

JUNIOR DATA SCIENTIST

June 2019 - April 2021

- Designed strategy for data acquisition, scraping and ingestion of the data periodically via pipeline for web sources.
- Developed and orchestrated ETL pipelines with Kubeflow for the ingested data.
- Participated, organised and communicated architectural and tool selection processes.
- Together with a colleague formulated the Scraping pipeline using Scrapy cloud.

LegalHorizon AG

Magdeburg, Germany

WORKING STUDENT

September 2018 - July 2019

- Developing tagging system for the large legal corpus, helps in refining search results.
- Scrapping of data from Eur-lex website
- Visualization of Topic models using PyLDAvis to better explain the search results.
- LHS_webapp, A RESTful API based web service for automation of tagging new documents in the database(SAS).
- Planning on implementing hierarchical deep learning for text classification and document tagging.

Bayer Pharmaceuticals AG

Berlin, Germany

DEEP LEARNING INTERN

September 2017 - January 2018

- Successfully developed proof of concept for application of Artificial Intelligence for single case processing.
- Implemented deep learning (LSTM) based NLP solution for unstructured text data.
- Worked on pre-processing of unstructured data, designing datapipeline and training a deep neural network.
- Gained experience on various natural language processing techniques such as POS tagging, Language Detection, Entity extraction and Clustering.
- Extracted data from PDFs, emails, scanned documents, etc and curated it using fuzzy string matching from manually annotated labels.

Ambuja Cements Ltd.

AUTOMATION ENGINEERING APPRENTICE

Kodinar, India

April 2015 - August 2015

- Developed Functional Block Diagrams for newly upgraded SCADA systems.
- Migration from Old DSC system to new SCADA system.
- Monitoring and Optimization of manufacturing process

Ambuja Cement Ltd.

AUTOMATION ENGINEERING INTERNSHIP

Kodinar, India

January 2015 - February 2015

- Learned the working of Cement Industry.
- Learned the essence of working in teams.

Skills

Programming Languages	Python, SQL, Shell Scripting
Deployment Frameworks	Docker, Kubernetes
Tools	ElasticSearch, DVC, MLFlow, Kubeflow, Streamlit
Libraries	Tensorflow, TFlearn, keras, numpy, pandas, scikit-learn, matplotlib, CNTK, spacy, nltk, fuzzywuzzy, Regex
Web Frameworks	Fast API, Flask
Platforms	Google Cloud Platform, AWS, Linux
Version Control	GIT
Scraping	Scrapy Framework, BeautifulSoup, Scrapinghub, crawlera
Web Automation	Selenium, Python Playwright
Languages	English (fluent), Gujarati (native), German (basic)

Projects

Intrusion Detection System using Microsoft Cognitive toolkit

Magdeburg, Germany

OTTO-VON-GUREICKE UNIVERSITY

April 2018 - July 2018

- This was continuation of our previous project, as our aim was to compare how differently the system behaves.
- This project explores the use of deep learning networks in Intrusion Detection Systems using NSL-KDD Cup 1999 dataset with Microsoft Cognitive toolkit.
- Explored three architectures, Auto-encoder, Convolutional Neural Networks and Recurrent Neural Networks after initial research.
- Our analysis showed that Convolutional Neural Networks performed slightly better than Auto-encoders and Recurrent Neural Networks

Implemented Intuition Detection System using Tensorflow

Magdeburg, Germany

OTTO-VON-GUREICKE UNIVERSITY

January 2017 - August 2017

- This project explored the use of deep learning networks in Intrusion Detection Systems using NSL-KDD Cup 1999 dataset using Tensorflow framework.
- Explored two architectures, Auto-encoder and Deep Belief Networks after initial research.
- Our analysis showed that Stacked Autoencoders with Greedy Layer-wise pre-training trained faster and achieved better results.

Autopyenv (An VSCode extension)

Würzburg, Germany

PERSONAL

- Developed an VSCode extension to manage different python environments.
- Useful in cases where developers working on different python projects. Makes it easy for them not to worry about creating and activating python virtual environments.
- <https://marketplace.visualstudio.com/items?itemName=jdvala.autopyenv>

FeelGoodNews

Würzburg, Germany

PERSONAL

- Built during Covid-19 times, as I was tired of reading negative news.
- Use pre-trained sentiment analysis model to filter news and present only positive news.
- Backend was done by me, with front-end by my friend and a freelancer
- 100+ downloads on both AppStore and PlayStore.

Other Mentions - small projects

PERSONAL

- python-lei - This project is wrapper for Leilex, legal entity identifier API. Includes ISIN-LEI conversion. Search LEI number using company name.
- zoom-audio-transcribe - Zoom Audio Transcription offline using vosk-api

Writing

Getting the basics right

[Blog](#)

FOUNDER & WRITER

May 2018 - PRESENT

- Drafted the progress of my Master's Thesis and issues I faced during each step.
- Notes with basics of Data Science and Machine Learning
- Introduction to Deep Learning, working with tensorflow
- Sometimes I write about things that are new to python.
- Generally, I write tech stuff.

Certificates

2018	Hadoop 101 , Cognitiveclass.ai	Germany
2017	Python 101 for Data Science , Cognitiveclass.ai	Germany
2017	Data Analysis with Python , Cognitiveclass.ai	Germany