

Prajnya Satish

+91 9845667465 | Bengaluru, India | prajnya.12@gmail.com | <https://www.linkedin.com/in/prajnyasatish>

SKILLS

Python: keras, gensim, nltk, tensorflow, huggingface, sklearn, pandas, flask, django

Machine Learning and AI: Classification, Feature engineering, Recurrent Neural Networks, LSTMs, Transformer Networks, LLMs

Programming Skills: Java, React, RESTful APIs, Javascript, Git, MySQL

Natural Language Processing: Question Answering, Word Sense Disambiguation, Coreference Resolution, Named Entity Recognition

WORK EXPERIENCE

Amelia, An IPSoft Company

New York, NY

R&D Engineer

Aug 2019 - Mar 2023

- Collaborated with Natural Language Understanding research team to improve core features of Amelia, a virtual conversational agent.
- Implemented and deployed models for specific tasks in language understanding such as SpanBERT based co-reference resolution for mention clustering within the project framework using gRPC and Protobuf. Improved coreference accuracy to 96.8%. Used Tensorflow, Keras, W&B and PyTorch for implementation and DL4J to serve and deploy models.
- Built and deployed customer facing inference based deep learning models for sentiment analysis, answer polarity, dialog acts, named entity recognition, etc. Models were based on BERT and other Transformer Networks. Took charge of the complete MLOps lifecycle from cleaning and maintaining data, training and deploying the models, and interfacing with external APIs such as chatGPT for client inference tasks.
- Incorporated knowledge graphs and lexical corpora such as OntoNotes and Verbnet for query understanding tasks.
- Added UI/UX features to the Amelia interface using React.js to make training, benchmarking and deployment of classifier models more convenient for cognitive engineers. This included building the RESTful APIs to allow users to make custom selections in the UI, using the selections to train models and displaying the final results to the user.
- Enhancing modules such as question answering, question generation, and open domain conversation among others.
- Improving the reasoning and entailment capabilities for domain-specific needs using Large Language Models.
- Built and deployed chatbot solutions using bot frameworks such as Rasa NLU, Amazon Lex, Amelia, etc

University of Colorado at Boulder

Boulder, CO

Student Assistant

Jan 2018 - May 2019

- Annotation: Performed manual annotation on the WSJ corpus to match verbs to their closest possible verb classes as part of the VerbNet sense annotation project.
- Developed and executed annotation scripts to convert Hindi-Urdu dataset with PropBank labels for visualization of dependency trees and phrase structure relations on the BRAT rapid annotation tool.
- Unified Verbs Index: Led development of the Unified Verbs Index (UVI) website - <https://uvi.colorado.edu/> using a full stack approach with Python, Flask, MongoDB, HTML, CSS, and Jinja Templating. UVI unifies knowledge base visualization of 4 different lexical corpora into a single website with over 5K monthly visitors.
- Collaborated with research teams during knowledge creation and website prototyping.

Leeds School of Business, University of Colorado

Boulder, CO

Student Assistant

Feb 2018 - May 2019

- Contributed to the development of a tool that converts unstructured optical data from PDFs and photographs into structured databases through text processing techniques. Worked on parsing Moody's manuals for grammar and semantics, achieving an accuracy rate of 86%.
- Information was picked from years' worth of Moody's manuals: the manuals were converted to text using OCR, parsed and corrected for grammar and semantics and then stored in structured databases.
- Tasks ran on a supercomputer in a cloud-based environment because of the exceedingly large size of the files and also the huge amount of computing power required. The processed data was later used to forecast future stock market trends based on past event.
- Used cloud-based supercomputers for running tasks due to large file sizes and computing power requirements. Successfully processed vast amounts of data from Moody's manuals stored in structured databases, which was later used for forecasting stock market trends accurately based on past events.

EmageVision Pte Ltd*Software Developer***Bangalore, India***May 2016 - Jul 2017*

- Robotics: Involved in a team that developed artificial intelligence capabilities on the NAO robot including its speech, movement, and vision.
- The aim was to deploy it in a self-sustained environment for customer support and engagement when support staff was unavailable.
- Built and deployed a question answering engine based on local knowledge base from scratch.
- Integrated Google Speech API for speech recognition.
- Integrated a chatbot using the open-source bot platform Rasa NLU - a project that builds machine learning models to capture the intent and keywords of the question.

PROJECTSgithub.com/PrajnyaSatish/Question-Answering-Engine**Bangalore, India****Question Answering Engine***Jan 2017 - Jul 2017*

- **Key Skills:** Python, SVM, Linear Regression, Question Answering, Search and Rank
- Developed and deployed a machine learning-based QA engine utilizing natural language processing techniques in Python, resulting in 90% accuracy for question classification.
- Leveraged feature engineering and trained SVMs and logistic regression models to accurately classify answer types.
- Designed and implemented a custom search engine to extract relevant documents from our knowledge base for use in answering user-generated questions, resulting in a 40% increase in successful query resolution.

github.com/PrajnyaSatish/MC-Task**Boulder, CO****Teaching Machines Reading Comprehension***Mar 2019 - May 2019*

- **Key Skills:** Python (Keras, Matplotlib, Pandas), Machine Comprehension, LSTMs, word embeddings.
- Implemented deep learning-based machine comprehension models using Keras, Matplotlib, and Pandas to achieve an accuracy score of 87% on a standardized reading comprehension test.
- Designed and executed experiments using multi-layer LSTM architectures with attention encoding to improve the model's ability to understand context in long-form text. Achieved a 25% increase in accuracy on questions related to complex paragraphs.
- Implemented multi-class classification to pick one of four correct options.

EDUCATION**University of Colorado at Boulder****Boulder, CO***MS in Computational Linguistics, Analytics, Search and Informatics**Graduation Date: May 2019***BNM Institute of Technology****Bangalore, India***BE in Telecommunications Engineering**Graduation Date: May 2015*