

Online Courses

- Machine Learning by Andrew Ng (Coursera)
- Deep Learning Specialization by deeplearning.AI (Coursera)
- Natural Language Processing by Christopher Manning (Youtube)
- Deep Learning for Natural Language Processing (Richard Socher)

Research Paper

“Home Automation using IoT and a Chatbot using Natural Language Processing”

IEEE Journal: 10.1109/IPACT.2017.8245185

Abstract:

Home automation - controlling the fans, lights and other electrical appliances in a house using Internet of things is widely preferred in recent days. In this paper, we propose a web application using which the fans, lights and other electrical appliances can be controlled over the Internet. The important features of the web application is that firstly, we have a chatbot algorithm such that the user can text information to control the functioning of the electrical appliances at home. The messages sent using the chatbot is processed using Natural Language processing techniques. Secondly, any device connected to the local area network of the house can control the devices and other appliances in the house. Thirdly, the web application used to enable home automation also has a security feature that only enables certain users to access the application. And finally, it also has a functionality of sending an email alert when intruder is detected using motion sensors.

Projects

Before 2020:

- Successfully completed an Emotionally Aware personal assistant using a Recurrent Neural Network.
- Successfully implemented a location disambiguation model for an app based company called Bundle (UK based).
- Built a learning tool for the intellectually challenged - Amogh Trust (NGO), resulted in winning JP Morgan's Chase Code for Good hackathon 2017, Bangalore.
- Implemented a document similarity model for recommendation of similar news articles for a news app.
- Built a rule-based chatbot from scratch using NLP for meetup planner app.

- Implemented (from scratch) a topic modeling algorithm for gaining personality insights from written speeches/essays.
- Data Visualizations of an online quiz event - Riddler.
- Web implementation of a sentiment analysis dashboard in collaboration with UC Berkeley.
- Made a prototype wearable in CAMTECH (Consortium of Affordable Medical Technologies) in organized in collaboration with MIT, Boston, MA

Work Experience

Company: Deep Learning Analytics (Now Opal AI)

Date: March 2020 to August 2021

Location: Remote

- **Action Detection from Meetings Audio and Emails:** Development and deployment of an application that extracts prosodic features, combines them with certain linguistic & structural features to identify action items. Helped the company automate their action item emails.
- **Question Detection from Emails:** Development and deployment of a CNN-based text classifier as an application that extracts linguistic features from emails and identifies questions from emails. Deployed using Tensorflow.js.
- **Video Summarization:** Development and deployment of an application which extracts prosodic features from audio and combines them with a recurrent neural network based text summation model to produce videos that summarize the results.
- **Question Answering on Video & Papers Corpus:** Re-calibrated and customized inference of the latest state-of-the-art BERT + Squad 2.0 Model to answer questions on a corpus of video transcriptions and technical paper abstracts that was successfully deployed.
- **Context-based text generation:** Customized and re-trained the OpenAI GPT-2 Language model on business articles for context-based text generation. Achieved perplexity score of 19 (better than the State-of-the-Art at the time on Reading comprehension(23)).
- **Follow-up Question Generation:** Created a transformer based model to generate follow-up questions given a context. Used OpenAI's GPT-2 model for language generation and Google Natural Questions dataset for training open-ended answer-question pairs.

- **Serendipity Engine** : Developed an application which scrapes LinkedIn for user profile data, feeds the data into a multi-pronged model which does feature-wise comparisons, generates scores and reasons for match. The application was used to match attendees of an event with similar backgrounds and goals.
- **Speaker Notes Generation**: Developed an application to generate speaker notes for a powerpoint presentation by generating text from slide content and by querying company knowledge base. Used the above search engine along with a BART-summarizer to generate notes.
- **Content and Media Search Engine**: Built a vector-based search engine for Accenture blogs, internal articles & podcasts using compressed T5 paragraph vectors and AWS elasticsearch with Question Answering capabilities.

Company: KPMG India - Lighthouse division

Date: July 2018 to March 2020

Location: Bangalore, India

- **Customer Review Analytics**: Development of an application that processed text reviews of customers for a major Japanese car manufacturer (350+ dealers in India), giving the manufacturer a complete understanding of their customer journey. (4-member team) Developed an application that performs data compilation, text cleaning, sentiment analysis, captures business touch-point, review context, and provides insightful phrases representing top promoters and detractors of a dealership. Developed a novel technique for phrase creation using a hybrid model that was successfully deployed to 350+ dealers.
- **Contract and Invoice Automation(Product)**: Developed a tool to automate contract and invoice processing that identifies important entities and notifies legal and financial teams of any discrepancies. The project reduced manual effort and time spent by the legal and financial teams. (3 member team) Developed a pipeline that includes image cleaning, localization, OCR, entity extraction, and Intelligent Rule Mapping of various business rules. The tool was sold to multiple clients including a large Indian power and mining conglomerate, a multinational beverage and brewing holdings company.
- **Assisted KPMG in winning multiple client engagements by developing effective Proof of Concept Softwares.**
- **Conducted a Series of Python and Machine Learning Workshops to help fellow colleagues get started with Programming and AI.**
- **Made multiple client proposals, presentations, and service catalogs.**

Company: Zycus, India

Date: July 2020 to July 2021

Location: Mumbai, India

- **Item Grouping for Supplier Filtering:** Developed a model with a supervised+unsupervised approach to group similar items based on their names & descriptions which helped in filtering suppliers for each item group based on price, location & other metadata. Used LSH for reducing comparisons, trained custom spacy NER & used Random Forest to classify based on NER+distance based features.

Accomplishments

- Winner of Mozilla Firefox Hackathon-prototyped music application.
- Made a prototype wearable in CAMTECH (Consortium of Affordable Medical Technologies) in organized in collaboration with MIT, Boston, MA
- VIT: Vice President, Computer Society of India. Helped organize events with more than 500 participants - 3 consecutive years
- Creation Labs: Actively involved in the management of a hackerspace. Team Captain of Autonomous Car Project.
- State level Chess player in the U-16 category.

Skills

Technical Skills: Natural Language Processing: Question Answering, Named Entity Recognition, Text summarization, Text Mining, Information Extraction, Text Classification, Text Generation, Clustering, Topic Modeling, Ranking, Search Engines , Chatbots (Intent Classification, Slot filling); Machine Learning (ML), Deep Learning, Image Processing, Audio Processing, Statistics, Data Cleaning, Data Analytics, Data Visualization, Regression, Data Engineering

Languages/Tools: Python , Java, HuggingFace , Spacy , NLTK , Numpy , Scikit-learn , Pandas , TensorFlow , Pytorch , AWS , Keras , Git , Github, Elasticsearch, SQL, NoSQL – MongoDB , Dialogflow , Django , PowerBI , C , C++ , HTML , JS