

Ambesh Shekhar

✉ ambesh.sinha@gmail.com ☎ +91-8825272821 🌐 AmbiTyga
🐙 Ambesh@Github 🔗 Ambesh@LinkedIn

EDUCATION

BIT MESRA

B.E. IN COMPUTER SCIENCE
AND ENGINEERING
Expected May 2020

COURSEWORK

UNDERGRADUATE

Fundamental of Data Structure
Object Oriented Programming
Advanced Design and Analysis of Algorithms
Artificial Intelligence
Operating Systems
Database System
Computer Networks
Computer Structure and Architecture
Software Engineering
System Programming
Compiler Designs and Principles
Computer Graphics and Design
Software Project Management

INDEPENDENT

Machine Learning Course (Coding Blocks)
CS 229 Machine Learning
CS 230 Deep Learning
Natural Language Processing Tensorflow (Coursera)
Improving Deep Learning Network(Coursera)
Sequence Model(Coursera)

SKILLS

PROGRAMMING

C/C++ • Python • Java
• R • Android • Dart
• Javascript

SCIENTIFIC LIBRARIES

Keras • Tensorflow • Pytorch •
Scikit-learn • Pandas • Numpy • NLTK
• Librosa • OpenCV

SOFTWARE AND TOOLS

PyCharm • Arduino • Raspbian • Flutter
• Android • MATLAB

PERSONALITY SKILLS

Leadership • Communication
• Management • Team Work

EXPERIENCE

Project Assistant | LABORATOIRE D'INFORMATIQUE DE GRENOBLE, FRANCE

JUNE 2020- JULY 2020

Working under Professor Genoveva Vargas-Solar in natural language processing field research. Providing optimal solution with my programming skills by extracting and cleaning data along with designing machine learning models to tackle the problems.

Research Assistant | BIRLA INSTITUTE OF TECHNOLOGY, MESRA

JAN 2020- JUNE 2020

Working under Professor Smita Pallavi on the applications of deep learning and computer science, and providing solutions for problems with my programming skills and knowledge in deep learning and machine learning.

RESEARCH WORK

🔗 Attention Based Consultation System for Medical Assistance

July 2020 - August 2020

The research aims to develop a dialogue system for medical assistance by training pre-trained attention based models over scraped data including informative texts and human-friendly interactions between patient and doctor. The research uses concept of encoder-decoder model to generate texts.

🔗 Reddit Mental Health Intervention Dialogue Dataset

July 2020 - August 2020

This dataset consists of conversation among peoples having mental illness. It has been extracted from Reddit, consisting more than 200k scraped posts. We perform analysis on the data to by generating human-understandable texts using generative pretrained networks.

🔗 COVID-19 Misinformation on Twitter: Multilingual Analysis

June 2020 - July 2020

The research leverages the fact-checking task using pre-trained deep learning networks fine-tuned over scraped data and manages to evaluate on multilingual data. The work shows demanding results over text data being shared through social medias, especially in checking for news related to Covid-19.

🔗 Character Level Pretrained Language Model for Extracting Support Phrases for Sentiment Labels

June 2020 - July 2020

The research aims to analyse texts from twitter by extracting phrases that describes the sentiment of a tweet. We developed a character-level pre-trained language model stacked over deep-learning networks for this task which could generate starting and ending index of the phrase at the character level.

🔗 QuesBELM: A BERT based Ensemble Language Model

March 2020 - May 2020

The research adapts the application of transfer-learning on text with ensemble learning by using attention-based models like BERT to answer complex problems which are mostly searched on Internet.

🔗 MemSem: A Multimodal framework for sentiment analysis

Feb 2020 - May 2020

The research shows analysis on user generated contents by using multi-modal deep learning methods on texts and images of a meme. It uses pretrained models to classify sentiments of the content.

LANGUAGE

- Native Hindi
- Advanced English
- Basic French

AREAS OF INTEREST

Natural Language Processing • Computer Vision • Speech Analysis • Machine Learning • Robotics • Algorithms

ACHIEVEMENTS

🏆 2nd Position in Internal Hackathon for SIH-2020

EXTRA CURRICULUM

🔗 Open-Source Contribution

</> Google-HashCode-2020

💡 Simple India Hackathon-2020

🤖 Techkriti-2019

⚡ e-Yantra Robotics competition-2018

🔧 Technoxian-2018

⚙️ Abhivyakti-2017

PROJECTS

🔍 Pothole Detection

Pothole Detection is full scale ML engine for real-time pothole detection. Working on the con-currency issue in the RCNN and increasing the accuracy of the output.

- Based on Masked-RCNN.
- Captures images using Raspberry-Pi and processes those images.
- Predicts pothole in the way by using trained model deployed using AWS sagemaker.

🔍 ASL-Classifier

A python script that utilizes the OpenCV and keras libraries to classify correct sign language

- Uses images of American Sign Languages and ConvNet architecture.
- Trained on preprocessed dataset and validated on self captured dataset.
- Uses methods of OpenCV to create user interface to test on real-time dataset.

🔍 hiLyted

A video highlights creator, clips video from the given input by performing acoustics analysis

- Downloads audio and video using youtube-dl.
- Uses Librosa to extract audio data and sample rate.
- Calculates and finds the right short time energy occurred in a 5 second window.
- Clips the video of that duration and concatenate all the video into a single one using MoviePy.
- Stores the highlighted video in local directory.

🔍 GifCaptioner

A deep learning network for GIFs captioning

- Downloads and trained on Tumblr gif dataset.
- Capture frames from a gif for deep learning operations.
- Normalises text for training.
- Uses VGG16 for frames classification and GloVe embeddings for captioning.
- Filters the text and store them in a text file.
- Use Bidirectional LSTM to handle the frames.

PUBLICATIONS

MemSem: A Multimodal framework for sentiment analysis, 37th International Conference for Machine Learning, 2020

QuesBELM: A BERT based Ensemble Language Model for Natural Questions, 5th IEEE International Conference on Computing, Communication and Security, 2020.