Abhijit Mahalunkar

+353-899893701 | abhijit65@gmail.com | www.abhijitm.com | Dublin,Ireland

Overview

Dedicated researcher specializing in machine learning, deep learning, and natural language processing. Proven expertise in advancing the state of the art in natural language processing through innovative research and development. Extensive experience in designing, implementing, and evaluating cutting-edge Large Language Models (LLM) architectures and fine-tuning pretrained LLMs. Collaborative team player committed to pushing the boundaries of LLMs and contributing to the broader field of artificial intelligence.

Education

Ph.D. Deep Learning, Technological University Dublin, Ireland. 2024
Thesis: *"The Complexity of Long-Distance Dependencies in Sequence Data and their Impact on the Representational Capacity and Performance of the Language Models"*, Supervised by Prof. John D. Kelleher

B.E. Electronics & Telecommunications, Goa College of Engineering, Farmagudi, Goa. 2010
Project: *"Design and Implementation of an Optimized Speech Recognition System"*
Computed speech signal data features, i.e., Cepstral Coefficients, and Mel-Frequency Cepstral Coefficients (MFCC), used Vector Quantization and Dynamic Time Warping (DTW) for matching.

Research Experience

Doctoral Research, Technological University Dublin, Ireland. 2017–2023
The research focused on improving foundational models, such as Large Language Models (LLMs), and assessing their proficiency in effectively capturing long-distance dependencies within datasets.

- Investigated the complexity of long-distance dependencies within sequence datasets, leading to the discovery that the dependencies within natural language decay with a broken power-law relationship.
- Confirmed that the occurrence of a broken power-law relationship of the dependency decay is attributed to the distribution of word co-occurrences in natural language.
- Investigated the representational capacity of language models, specifically exploring Recurrent and Transformer-based architectures. Utilized datasets of artificial grammars to analyze their performance.
- Optimized the hyper-parameters of language models through an enhanced grid-search process by refining the hyper-parameter space, leading to improved performance of the language model.
- Explored the scaling properties of natural languages and their intricate relationship with the representational capacity of language models, providing valuable insights for improving language models.
- Conducted training of language models and fine-tuned pre-trained LLMs, exploring their capacity to effectively model Long-Distance Dependencies. Prepared datasets for training and fine-tuning.
- Explored the predictability of sequences and their interconnection with the compression capabilities of language models, contributing valuable insights to the explainability of language models.

Work Experience

Technical Consultant, Qubiseed Technologies LLP, Goa, India. Nov 2015–Dec 2017

- Developed a differential diagnosis system for endocrine diseases by consulting medical experts.
- Developed the web application architecture for the doctor's appointment system using Amazon AWS.

Co-founder & Product Architect, Spitiq, Goa, India. April 2015–July 2017

- Led and managed a team that designed wireless sensor nodes using Atmega128RFA1 microcontroller, smart home sensors, and electric switching to be integrated with a home automation system.
- Implemented activity discovery for the home automation system by processing the interleaved sensor data collected via the wireless sensor network to assist in optimal decision-making.

Technical Consultant, SmartKlock Inc., Austin, TX. Oct 2014–Oct 2015

- Led a team that developed a social media device utilizing the BeagleBone Black operating on the Android OS, seamlessly combining social media functionality with that of a table clock.

Project Assistant, National Institute of Oceanography (NIO), Goa, India. Nov 2010–Sept 2014

- A member of the team engaged in the development and deployment of the Autonomous Underwater Vehicle (AUV-MAYA) and Autonomous Vertical Profiler (AVP) for scientific ocean data collection.
- Developed a Hardware-In-Loop Simulator for AUV-MAYA to minimize the need for extensive field trials by providing a controlled laboratory environment for testing the functionality of AUV-MAYA.

- Developed a communication protocol to transmit the data between the AVP and NIO server via the Iridium satellite constellation, enabling the AVP to operate as a Lagrangian drifter.

Machine Learning Projects

Fine-Tuned a Transformer-Based LLM on Named Entity Recognition (NER) task 2023

- Fine-tuned pre-trained BERT, DistilBERT, and XLNet LLMs on the MultiNERD dataset (Multilingual, Multi-Genre, and Fine-Grained) for the NER task. Compared the performance of the fine-tuned LLMs.

Data-Driven Predictive Modeling for Housing Price Trends 2023

- Conducted extensive data collection of housing data, acquiring diverse features influencing house prices. Applied advanced feature engineering techniques to eliminate correlations and identify pivotal features contributing to enhanced predictive accuracy in housing price models.
- Implemented time-series regression models, e.g., ARIMA, LSTM, and SVM, to predict housing price trends for diverse regions. Leveraged statistical techniques and data preprocessing to extract meaningful insights and optimize predictive models for robust performance in varying real estate markets.

Advancing Human Mobility Predictability 2020

- Identified flaws in the computation of the upper bound of human mobility predictability, highlighting the impact of assumptions, specifically the assumed Markovian nature of human mobility.
- Demonstrated through statistical tests on real-world mobility datasets that human mobility exhibits scale-invariant long-distance dependencies, underscoring the need to account for long-distance dependencies for accurate predictability assessments.

Innovative Image Captioning Model Development 2018

- Part of the team that developed an image captioning model that enhances caption diversity and specificity by incorporating unsupervised training with a learning signal from an Image Retrieval model.
- Improved image captioning in generating more diverse and novel captions for similar images.

Teaching

Senior Demonstrator, Technological University Dublin, Ireland Sept 2017–June 2021

- Conducted labs and tutored students in Deep Learning, Machine Learning, and Databases

Instructor, CTYI - Dublin City University, Ireland June–July 2019

- Designed and delivered a course on robotics and embedded systems for high school students.

Technical Skills

Python, Java, C, C++, R, HTML/CSS, JavaScript, PHP, SQL, MATLAB, Octave, PyTorch, TensorFlow, Keras, Scikit-Learn, OpenCV, Eclipse, Visual Studio, Android SDK, Django, Ruby on Rails, Node.JS.

Awards & Grants

TU Dublin Scholarship to pursue a Ph.D. at Technological University Dublin, Ireland. 2019
ADAPT Auxiliary Fund to enhance computational capabilities for Language Model training. 2019
DIT Fiosraigh Award, to pursue an MPhil at Dublin Institute of Technology, Ireland. 2017
NVIDIA GPU grant of NVIDIA TITAN Xp GPU to enhance the computational capability. 2017

Selected Publications

[1] A. Mahalunkar, J. D. Kelleher (2020). Mutual Information Decay Curves and Hyper-parameter Grid Search Design for Recurrent Neural Architectures. *The 27th International Conference on Neural Information Processing, ICONIP 2020.* doi: https://doi.org/10.1007/978-3-030-63823-8_70

[2] A. Mahalunkar, J. D. Kelleher (2019). Multi-Element Long Distance Dependencies: Using SPk Languages to Explore the Characteristics of Long-Distance Dependencies. *ACL: The Workshop on Deep Learning and Formal Languages: Building Bridges.* doi: https://aclanthology.org/W19-3904/

[3] V. Kulkarni, A. Mahalunkar, B. Garbinato, J. D. Kelleher (2019). Examining the Limits of Predictability of Human Mobility. *Entropy.* doi: https://www.mdpi.com/1099-4300/21/4/432

[4] V. Kulkarni, A. Mahalunkar, B. Garbinato, J. D. Kelleher (2019). On the Inability of Markov Models to Capture Criticality in Human Mobility. *Artificial Neural Networks and Machine Learning - ICANN 2019: Image Processing.* doi: https://doi.org/10.1007/978-3-030-30508-6_39

[5] A. Mahalunkar, J. D. Kelleher (2018). Using Regular Languages to Explore the Representational Capacity of Recurrent Neural Architectures. *Artificial Neural Networks and Machine Learning - ICANN 2018.* doi: https://doi.org/10.1007/978-3-030-01424-7_19