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
Relevant Courses: Machine Learning, Deep Learning, and Algorithms & Approximation.

B.E. Electronics & Telecommunications, Goa College of Engineering, Farmagudi, Goa.   **2010**
Project: *"Design and Implementation of an Optimized Speech Recognition System"*
Relevant Courses: Probability Theory and Random Processes, Artificial Neural Networks, Robotics, Digital Signal Processing, Data Communication, and Control System Engineering.

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.

- Discovered that dependencies within natural language decay with a broken power-law relationship, shedding light on the complexity of long-distance dependencies within sequence datasets.
- 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.
- Analyzed the performance of large language models, including Recurrent and Transformer-based architectures, using datasets of artificial grammars to explore their representational capacity.
- Enhanced the performance of neural architectures by optimizing hyper-parameters through a refined grid-search process, resulting in improved model performance.
- Provided valuable insights into the scaling properties of natural languages (e.g., Zipf's, Taylor's, and Ebeling's Law) and their intricate relationship with language models' representational capacity, contributing to advancements in large language model development.
- 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.
- Trained and fine-tuned language models, including pre-trained LLMs, to effectively model Long-Distance Dependencies, contributing to understanding large language model capabilities and limitations. Prepared datasets for training and fine-tuning to ensure model accuracy and reliability.

WORK EXPERIENCE

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

- Collaborated with medical experts to develop a differential diagnosis system for endocrine diseases, leveraging domain knowledge and technical expertise to create an accurate and reliable diagnostic tool.
- Demonstrated full-stack development proficiency by architecting a resilient doctor appointment system hosted on Amazon AWS, optimizing scalability and performance to streamline appointment scheduling processes and enhance overall efficiency.
- Engineered a robust backend infrastructure on Amazon AWS, leveraging services like Lambda, API Gateway, and DynamoDB to ensure the system's scalability, security, and optimal performance.

Product Architect, Spitiq, Goa, India.   **April 2015–July 2017**

- Spearheaded and supervised a team in developing wireless sensor nodes, paving the way for seamless integration into a comprehensive home automation framework.
- Engineered an innovative temporal activity detection system using recurrent neural networks within the home automation ecosystem, processing intricate sensor data collected from the wireless sensor network to discern and facilitate optimal decision-making processes.

Technical Consultant, SmartKlock Inc., Austin, TX.   **Oct 2014–Oct 2015**
- Led and guided a team in the creation of a social media device, ingeniously merging social media features with the functionality of a table clock, offering users a novel and versatile digital experience.

Project Assistant, National Institute of Oceanography (NIO), Goa, India.   **Nov 2010–Sept 2014**
- Played a pivotal role in developing Autonomous Underwater Vehicles (AUV-MAYA and AVP), instrumental in advancing scientific ocean data collection capabilities.
- Engineered a sophisticated Hardware-In-Loop Simulator for AUV-MAYAn using Python, reducing the necessity for extensive field trials by establishing a controlled laboratory environment to rigorously test the functionality of AUV-MAYA under various conditions.
- Innovated and implemented a robust communication protocol to facilitate seamless data transmission between the AVP and NIO server via the Iridium satellite constellation, enabling the AVP to operate effectively as a Lagrangian drifter.
- Orchestrated the setup and construction of an in-house Linux server infrastructure, leveraging advanced technologies and best practices to ensure optimal performance and reliability.
- Architected and implemented a highly sophisticated backend system tailored for seamless communication, effectively facilitating data transmission and processing across various system components.
- Designed and deployed a cutting-edge web application featuring dynamic visualization capabilities, empowering stakeholders to extract actionable insights from the data collected by Autonomous Vertical Profilers (AVPs) and enhancing overall system efficiency and effectiveness.

MACHINE LEARNING PROJECTS

Fine-Tuned a Transformer-Based LLM on Named Entity Recognition (NER) task   **2023**
- Conducted fine-tuning of large language models (LLMs), including BERT, DistilBERT, and XLNet, utilizing the MultiNERD dataset, a Multilingual, Multi-Genre, and Fine-Grained dataset tailored for Named Entity Recognition (NER) and Disambiguation tasks, leveraging the HuggingFace libraries.
- Systematically evaluated the performance of the fine-tuned large language models on the NER task, providing comprehensive insights into their effectiveness and suitability for real-world applications.

Data-Driven Predictive Modeling for Housing Price Trends   **2023**
- Spearheaded the collection of extensive housing data, meticulously acquiring diverse features influencing house prices. Applied advanced feature engineering techniques to eliminate correlated features and identify pivotal features contributing to heightened predictive accuracy in housing price models.
- Led the implementation of innovative methods aimed at decoupling distinct housing price trends across diverse market segments, resulting in an enhancement in the accuracy of housing price forecasting.
- Implemented cutting-edge time-series regression models including ARIMA, LSTM, and SVM, to forecast housing price trends across diverse regions.
- Leveraged expertise in data analysis and predictive modeling to design tailored approaches for dissecting complex housing market dynamics, improving the accuracy of housing price predictions.

Advancing Human Mobility Predictability   **2020**
- Identified critical flaws in the methodology used to compute the upper bound of human mobility predictability, exposing the significant impact of underlying assumptions, particularly the assumption of the Markovian nature of human mobility.
- Conducted statistical analyses on real-world mobility datasets, revealing compelling evidence that human mobility exhibits scale-invariant long-distance dependencies. Emphasized the importance of considering long-distance dependencies for accurate predictability assessments in studies on mobility.

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.

Design and Implementation of an Optimized Speech Recognition System   **2010**
- Applied expertise in signal processing algorithms to optimize feature extraction methods, enhancing the effectiveness and reliability of speech signal analysis for diverse applications.

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

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**

GRANTS

Travel Grant by Naver Labs	**Aug 2019**

To attend the ACL 2019 workshop on Deep Learning and Formal Languages: Building Bridges.

ENNS Student Travel Grant	**Oct 2018**

To attend the International Conference on Artificial Neural Networks, Rhodes, Greece.

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