

# Abhijit Mahalunkar

+91-9764594934 | abhijit65@gmail.com | www.abhijitm.com | Goa, India

OVERVIEW	I am a skilled researcher with expertise in training and fine-tuning Language Models and modeling sequential data. My research work, complemented by hands-on experience in machine learning, positions me as a dynamic professional, bridging the gap between theoretical knowledge and practical applications.	
EDUCATION	<b>Ph.D. Deep Learning</b> , Technological University Dublin, Ireland.	(exp.) 2024
	<i>"The Complexity of Long-Distance Dependencies in Sequence Data and their Impact on the Representational Capacity and Performance of the Language Models"</i> , Supervised by Prof. John D. Kelleher	
	<b>B.E. Electronics &amp; Telecommunications</b> , Goa College of Engineering, Farmagudi, Goa.	2010
SKILLS	<b>Research:</b> Data Collection, Statistical Analysis, and Research Design and Methodology. <b>Large Language Models:</b> Trained and fine-tuned language models, e.g., BERT, GPT, LLaMa, PaLM, and Mixtral (using HuggingFace libraries) on downstream tasks. <b>Data Processing:</b> Examined the scaling properties of large natural language datasets. <b>Programming Languages:</b> C, C++, Visual C++, Python, Java, R, SQL, LabView, MATLAB, Octave. <b>Libraries:</b> PyTorch, TensorFlow, Keras, Scikit-Learn, Pandas, Matplotlib, OpenCV, CUDA. <b>Tools:</b> Anaconda, GIT, Jupyter Notebook, Eclipse, Visual Studio, Android SDK, Nokia Qt Applications.	
RESEARCH EXPERIENCE	<b>Doctoral Research</b> , Technological University Dublin, Ireland.	2017–2023
	The research is centered on uncovering the limitations of Language Models (both Transformer-based and RNN-based) in effectively capturing Long-Distance Dependencies present in datasets.	
	<ul style="list-style-type: none"><li>Researched to investigate Long-Distance Dependencies within large sequence datasets, identifying variations in the dependency decay patterns across different sequence datasets. Notably, discovered a broken power-law relationship in the dependency decay of natural language datasets.</li><li>Analyzed the broken power-law dependency decay in natural language datasets, elucidating this phenomenon through an investigation of word co-occurrences within these datasets.</li><li>Compared the representational capacity of RNN-based and Transformer-based Language models using artificial grammars, providing valuable insights into their relative strengths and weaknesses.</li><li>Accelerated the grid-search optimization process for Language Model hyper-parameters by narrowing down the hyper-parameter space, informed by examining how different hyper-parameters are influenced by the dependency decay pattern, resulting in an improved and finely tuned model performance.</li><li>Investigated the nuanced relationship between the scaling properties of natural language datasets and the evaluation of Language Models, offering rich insights for the enhancement of Language Models.</li><li>Trained language models (RNN-based and Transformer-based) using TensorFlow and PyTorch. Also, fine-tuned pre-trained Large Language Models (transformer-based) and investigated their ability to learn Long-Distance Dependencies.</li></ul>	
WORK EXPERIENCE	<b>Technical Consultant</b> , Qubiseed Technologies LLP, Goa, India.	Nov 2015–Dec 2017
	<ul style="list-style-type: none"><li>Developed a differential diagnosis system for endocrine diseases by consulting medical experts.</li><li>Developed the web application architecture for the doctor's appointment system using Amazon AWS.</li></ul>	
	<b>Co-founder &amp; Product Architect</b> , Spitiq, Goa, India.	April 2015–July 2017
	<ul style="list-style-type: none"><li>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.</li><li>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.</li></ul>	
	<b>Technical Consultant</b> , SmartKlock Inc., Austin, TX.	Oct 2014–Oct 2015
	<ul style="list-style-type: none"><li>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.</li></ul>	
	<b>Project Assistant</b> , National Institute of Oceanography (NIO), Goa, India.	Nov 2010–Sept 2014
	<ul style="list-style-type: none"><li>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.</li></ul>	

	<ul style="list-style-type: none"> <li>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.</li> <li>Developed a communication protocol to transmit the data between AVP and NIO server via the Iridium satellite constellation and a web app to display the real-time location of AVP and the collected data. This enabled the AVP to operate as a Lagrangian drifter, gathering ocean column data amidst the demanding conditions of the monsoon season over three months.</li> </ul>	
MACHINE LEARNING PROJECTS	<b>Fine-Tuned a Transformer-Based LLM on Named Entity Recognition task</b> <ul style="list-style-type: none"> <li>Dataset used: MultiNERD: A Multilingual, Multi-Genre, and Fine-Grained Dataset for Named Entity Recognition (and Disambiguation).</li> <li>Fine-tuned pre-trained BERT, DistilBERT, and XLNet LLMs on the MultiNERD dataset for Named Entity Recognition task using HuggingFace libraries.</li> <li>Compared the performance of the fine-tuned LLMs on the Named Entity Recognition task.</li> </ul>	2023
	<b>Design of Speech Synthesis System</b> <ul style="list-style-type: none"> <li>Extracted speech parameters, i.e., Mel-Frequency Cepstral Coefficients (MFCC) from phonemes, created a database for text-to-phoneme lookup, and synthesized speech using TD-PSOLA.</li> </ul>	2011
	<b>Design and Implementation of an Optimized Speech Recognition System</b> <ul style="list-style-type: none"> <li>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.</li> </ul>	2010
TEACHING	<b>Senior Demonstrator</b> , Technological University Dublin, Ireland <ul style="list-style-type: none"> <li>Conducted labs and tutored students in Deep Learning, Machine Learning, and Databases</li> </ul> <b>Instructor</b> , CTYI - Dublin City University, Ireland <ul style="list-style-type: none"> <li>Designed and delivered a course on robotics and embedded systems for high school students.</li> </ul>	Sept 2017–June 2021 June–July 2019
AWARDS & GRANTS	<b>TU Dublin Scholarship</b> to pursue a Ph.D. at Technological University Dublin, Ireland. <b>ADAPT Auxiliary Fund</b> to enhance computational capabilities for Language Model training. <b>DIT Fiosraigh Award</b> , to pursue an MPhil at Dublin Institute of Technology, Ireland. <b>NVIDIA GPU grant</b> of NVIDIA TITAN Xp GPU to enhance the computational capability.	2019 2019 2017 2017
PUBLICATIONS	[1] A. Mahalunkar, J. D. Kelleher (2020). Mutual Information Decay Curves and Hyper-parameter Grid Search Design for Recurrent Neural Architectures. <i>The 27th International Conference on Neural Information Processing, ICONIP 2020</i> . doi: <a href="https://doi.org/10.1007/978-3-030-63823-8_70">https://doi.org/10.1007/978-3-030-63823-8_70</a> [2] A. Mahalunkar, J. D. Kelleher (2019). Multi-Element Long Distance Dependencies: Using SPk Languages to Explore the Characteristics of Long-Distance Dependencies. <i>The Workshop on Deep Learning and Formal Languages: Building Bridges</i> . doi: <a href="https://aclanthology.org/W19-3904/">https://aclanthology.org/W19-3904/</a> [3] V. Kulkarni, A. Mahalunkar, B. Garbinato, J. D. Kelleher (2019). Examining the Limits of Predictability of Human Mobility. <i>Entropy</i> . doi: <a href="https://www.mdpi.com/1099-4300/21/4/432">https://www.mdpi.com/1099-4300/21/4/432</a> [4] V. Kulkarni, A. Mahalunkar, B. Garbinato, J. D. Kelleher (2019). On the Inability of Markov Models to Capture Criticality in Human Mobility. <i>Artificial Neural Networks and Machine Learning - ICANN 2019: Image Processing</i> . doi: <a href="https://doi.org/10.1007/978-3-030-30508-6_39">https://doi.org/10.1007/978-3-030-30508-6_39</a> [5] A. Mahalunkar, J. D. Kelleher (2018). Using Regular Languages to Explore the Representational Capacity of Recurrent Neural Architectures. <i>Artificial Neural Networks and Machine Learning - ICANN 2018</i> . doi: <a href="https://doi.org/10.1007/978-3-030-01424-7_19">https://doi.org/10.1007/978-3-030-01424-7_19</a> [6] A. Mahalunkar, J. D. Kelleher (2018). Understanding Recurrent Neural Architectures by Analyzing and Synthesizing Long Distance Dependencies in Benchmark Sequential Datasets. <i>arXiv e-prints</i> .	2019 2019 2017 2017 2018 2018
REFERENCES	<b>Prof. John D. Kelleher</b> Professor of Computer Science Hamilton Institute, Maynooth University, Ireland <a href="mailto:john.kelleher@mu.ie">john.kelleher@mu.ie</a>	<b>Chetan Desai</b> Co-Founder & CEO Qubiseed Technologies LLP, India <a href="mailto:chetan@qubiseed.com">chetan@qubiseed.com</a>