

Arjun Manoharan

+91-9840256209 • arjunmanoharan2811@gmail.com

Education

Program	Institution	%/CGPA	Year of completion
M.S. (Computer Science and Engg.)	IIT Madras	8.8	2020
B.E. (Electronics and Comm. Engg.)	Meenakshi Sundararajan, Chennai	8.45	2015

Technical Skills

Languages: Python **Framework/Technical Libraries:** Pytorch, Tensorflow, Keras, Flask, Dash, NLTK, spacy, MLflow, LangChain, Huggingface, streamlit, wandb

Industrial Work Experience

Verisk Analytics

Data Scientist II

July 2020 - Present

Kondapur, Hyderabad

○ Natural Language Processing and Reinforcement Learning

- Worked on Reinforcement Learning based extractive-abstractive summarization model.
- Co-authored a paper explaining the effect of intermediate pretraining of extractive summarization models for scientific documents. The paper was accepted at SDP workshop ACL.
- Retaining salient sections of an article based on NER, OpenIE. Created a demo using flask hosting 8 methods.
- Worked on enhancing the factual consistency of abstractive summarization systems. Integrated dependency graph structure in summarization models. Experimented on CNN/DM and Xsum datasets.
- Built a RAG system using LLAMA2, Qdrant to identify the business type given the url of the business.

○ Computer Vision

- Built image retrieval and image classification systems using CLIP.
- Created a pipeline consisting of a classifier and transformer-based model for detecting important entities from images of business card. Created an endpoint using MLFlow.
- Working on transformer-based image captioning and a visual question-answering(VQA) system to detect hazards in a given image.

○ Pandemic Modelling

- Created a dashboard using dash for modelling the number of people infected due to COVID-19.
- Built a model to predict the Degree and Epsilon for the given configuration of the NPIs.

Verizon Data Services India

Analyst

June 2015 - July 2017

Guindy, Chennai

- Was part of the Customer Relationship Management Manager(CRMM) a revenue critical team. Specifically, I was working on the qualification engine that is used to qualify products for all wireline customers of Verizon.

Key Projects

Ask Me Anything!!!

Aug 2023 - Present

Verisk Analytics

- Created a chatbot to answer questions about a document. Used LLAMA2, Pinecone and sentence transformer.
- Used streamlit for frontend.

Tags: LLM, Pinecone, Llama2, streamlit

Crowdsourcing - RL for task assignment

Aug 2020 - Nov 2020

Verisk Analytics

- The objective of this project is to build a Reinforcement Learning agent that is able to effectively delegate labelling tasks to crowd workers.
- Built a RL agent that takes in the current set of tasks and workers as inputs for task delegation. The RL agent has time varying action space.

Tags: Reinforcement Learning for task delegation

RL on the Edge

Guide: Prof. Pratyush Kumar Panda, Prof. Balaraman Ravindran

Aug 2018 - Dec 2018

IIT Madras

- This project tried to examine how quantization tricks work on RL agents. Surprisingly for games like Pong on Atari, were able to see a 75 % reduction in bandwidth.
- RL can be used on edge devices, especially on mobile games. In future, there is a possibility to deploy RL agents which adapt according to the way a player plays the game.

Tags: Edge devices, Reinforcement Learning for mobile games, Quantization

Internship

University of Texas at Dallas

Research Intern

July 2019 - Sept 2019

STARLING Lab

- Worked with Prof.Sriraam Natarajan, Prof.Balaraman Ravindran and Prof.Prasad Tadepalli in combining planning and reinforcement learning.
- Presented papers at ICAPS 2021 and StarAI 2021.

National Institute of Ocean Technology

Intern

Feb 2015 - May 2015

- Worked on the project "Wireless Sensor Network Based Spatial Parameters Measurement For Ocean Applications" under the guidance of Mr.Shijo Zacharia.

Honours

- Awarded **Endowment Prize** for excelling in engineering graphics.
- Awarded **SpotLight Award** at Verizon Data Service India for actively participating in the design and development of a project from scratch.
- Awarded **Travel Grant** by Robert Bosch Center for Data Science and Artificial Intelligence, IIT Madras for the proposal Hierarchical Relational Reinforcement Learning.

Publications

1. **Effect of Pretraining on Extractive Summarization for Scientific Documents** Yash Gupta, Pawan Sasanka Ammanamanchi, Shikha Bordia, **Arjun Manoharan**, Deepak Mittal, Ramakanth Pasunuru, Manish Shrivastava, Maneesh Singh, Mohit Bansal, Preethi Jyothi - [SDP Workshop ACL 2021]
2. **Deep RePReL-Combining Planning and Deep RL for acting in relational domains**-Harsha Kokel, **Arjun Manoharan**, Sriraam Natarajan, Balaraman Ravindranm Prasad Tadepalli - [Deep RL Workshop NeurIPS 2021]
3. **Dynamic probabilistic logic models for effective abstractions in RL**-Harsha Kokel, **Arjun Manoharan**, Sriraam Natarajan, Balaraman Ravindranm Prasad Tadepalli-[StarAI 2021]
4. **RePReL: Integrating Relational Planning and Reinforcement Learning for Effective Abstraction** -Harsha Kokel, **Arjun Manoharan**, Sriraam Natarajan, Balaraman Ravindranm, Prasad Tadepalli-[ICAPS 2021]
5. **Option Encoder: A Framework for Discovering a Policy Basis in Reinforcement Learning** **Arjun Manoharan**, Rahul Ramesh, Balaraman Ravindran[ECML2020]

Positions of Responsibility

1. Teaching Assistant, Reinforcement Learning

CS6700, Prof. Balaraman Ravindran

Aug 2018 - Nov 2018

IIT Madras

2. Teaching Assistant, Reinforcement Learning

ACM Summer School on Machine Learning and Data Science

April 2018

Goa University

Conceptualized and handled lab sessions on Reinforcement Learning.

3. Teaching Assistant, Introduction to Programming Lab

CS1100, Prof.Mithesh Khapra

Aug 2017 - May 2018

IIT Madras

Course Work

Core Courses: Reinforcement Learning, Deep Learning, Pattern Recognition, Advanced Data Structures and Algorithms

Elective: Linear Algebra and Random Processes, Topics in Reinforcement Learning