

# Chathuri Jayaweera

[chathuri.jayawee@ufl.edu](mailto:chathuri.jayawee@ufl.edu) | [linkedin.com/chathurijayaweera](https://www.linkedin.com/in/chathurijayaweera) | [chestnutjay.github.io](https://github.com/chestnutjay)

## INTERESTS

---

- ◇ Natural Language Processing ◇ Natural Language Understanding & Inference
- ◇ Data Science & Machine Learning ◇ Explainable AI ◇ Commonsense Reasoning
- ◇ LLM – based Evaluation

## EDUCATION

---

### University of Florida

Ph.D. in Computer Science, Advisor: Prof. Bonnie Dorr  
Department of Computer and Information Science and Engineering  
Advanced to Candidacy January 2026; Expected Graduation August, 2026.

*Gainesville, Florida*

2022-Present

### University of Moratuwa

Honours Degree of Bachelor of Science in Engineering  
Specialisation: Electronic and Telecommunication Engineering  
Dean's Honour List in Semester 1, 2, 6, 7 and 8.

*Moratuwa, Sri Lanka*

2015-2019

## PUBLICATIONS

---

### Peer-reviewed Publications

- Chathuri Jayaweera, Bonnie J. Dorr. [From Disagreement to Understanding: The Case for Ambiguity Detection in NLI](#). In Proceedings of the 4th Workshop on Perspectivist Approaches to NLP (NLPerspectives), EMNLP 2025
- Chathuri Jayaweera, Sangpil Youm, and Bonnie J. Dorr. [AMREx: AMR for Explainable Fact Verification](#). In Proceedings of the 7th Fact Extraction and VERification Workshop (FEVER), EMLNP 2024
- Sangpil Youm, Brodie Mather, Chathuri Jayaweera, Juliana Prada and Bonnie J. Dorr. [DAHRS: Divergence-Aware Hallucination Remediated SRL Projection](#). In International Conference on Applications of Natural Language to Information Systems, Springer, 2024.
- Manuel Nunez Martinez, Sonja Schmer-Galunder, Zoey Liu, Sangpil Youm, Chathuri Jayaweera, and Bonnie J. Dorr. [Balancing Transparency and Accuracy: A Comparative Analysis of Rule-Based and Deep Learning Models in Political Bias Classification](#). In Proceedings of the 2nd Workshop on Social Influence in Conversations (SICon 2024), EMNLP 2024.
- Chathuri Jayaweera and Gihan Dias. [Automatic Classification, Visualization and Analysis of Errors in Machine Translation](#). In Extended Abstracts of International Conference on Information and Automation for Sustainability (ICIAFS), IEEE, 2021.
- Sutharsan, Mahendren, Shehan Rajakaruna, S. Y. Jayaweera, Chathuri Jayaweera, and Subramaniam Thayaparan. [Vision-Based Adaptive Traffic Light Controller for Single Intersection](#). In proceedings of the International Conference on Information Technology Research (ICITR), IEEE, 2020.

### Manuscripts Under Review

- Chathuri Jayaweera, Brianna Yanqui, Bonnie J. Dorr. [Filling the Gap: Is Commonsense Knowledge Generation useful for Natural Language Inference?](#). Annual Meeting of the Association for Computational Linguistics (ACL). Under-review. arXiv: 2507.15100, 2026
- Kweku Yamoah, Godfred Agyapong, Neel Parekh, Detravious Brinkley, Chathuri Jayaweera, Shlok Gilda, Bonnie Dorr, Emmanuel Dorley and Kevin Scroggins. [An Elicitation-Matrix Approach to Pragmatic Context Modeling in Low-Resource Machine Translation: The Case of Akuapem Twi](#). FLAIRS. Accepted.

## Preprints

- Chathuri Jayaweera and Bonnie J. Dorr. [BLADE: Better Language Answers through Dialogue and Explanations](#). Preprint 2026

## WORK HISTORY

---

Aug 2022 Present	Graduate Research Assistant (Research grant: DARPA INCAS (2023-2025)) <i>The Natural Language Research &amp; Culture (NLP&amp;C) Laboratory, CISE, University of Florida</i> Research ambiguity detection and resolution in Natural Language Inference (NLI), targeted entity extraction for sentiment and stance detection, and the development of explainable systems using foundation models, semantic and dependency representations, and graph neural networks (GNNs).
Aug 2022 Present	Graduate Teaching Assistant <i>CAP4641, CAI6307 - Natural Language Processing, CISE, University of Florida</i> Assist undergraduate students in course projects. Mentor graduate student teams in conducting NLP research, guide them in designing experimental pipelines, and collaborate on publications.
Sep 2021 Jun 2022	Cognius.ai - AI Research Engineer Conducted research on few-shot intent classification in conversational AI, Fine-tuned language models for few-shot learning with minimal data, and generated synthetic data to improve intent classification efficiency.
Feb 2020 Sep 2021	Graduate Research Assistant (Research grant: AHEAD) <i>Natural Language Processing Center (NLPC), CSE, University of Moratuwa</i> Conducted research on enhancing low-resource language translation (Sinhala-English) performance through post-processing techniques by identifying and addressing frequent error types in the translation process.
Jun 2018 Dec 2018	Forestpin (Pvt.) Ltd. - Machine Learning Research Intern Applied statistical modeling techniques to develop sales forecasting models and implemented inventory prediction algorithms, optimizing management and decision-making in supermarket outlets.

## SELECTED PROJECTS

---

### **BLADE: Better Language Answers through Dialogue and Explanations (2025-2026)**

Conducted research study on the impact of using a knowledge-grounded RAG-based conversational assistant on student performance in an undergraduate-level course. Established the technical foundation and intellectual framework for an NSF proposal centered on BLADE

Stack: RAG-based conversation assistant, Python, Linear Mixed-Effect Modeling, Statsmodels

### **LLMs as Commonsense Knowledge generators for Natural Language Inference (NLI) (2025)**

Evaluated LLMs as commonsense knowledge sources by generating axioms for NLI predictions on SNLI and ANLI datasets. Implemented an LLM-as-judge framework to scale evaluation and assess axiom factuality.

Stack: Llama-3.1-70B-Instruct, gpt-oss-120b, Claude-3.5-Sonnet, Python

### **Fact Verification via AMR with Explainability (2024)**

Developed an explainable fact verification pipeline using AMR graphs to assess meaning containment between claim and evidence. Combined AMR evaluation metrics with cosine similarity and provided post-hoc graph alignments for prediction interpretability.

Stack: Python, IBM transitioned-based AMR parser, Smatch,  $S^2$ match, SentenceBERT

### **Targeted Sentiment Analysis through Convolutional Networks (2023)**

Developed a GCN-based entity recognition system combining dependency parsing, rule-based features, and token embeddings for node classification. Integrated the model into a targeted sentiment analysis pipeline and evaluated on SemEval 2014 ABSA dataset (DARPA INCAS).

Stack: Python, SpaCy, Dependency trees, Graph Convolutional Networks, DistilBERT

### **Automatic Error Visualizing and Analyzing Tool (2021)**

Developed a tool for the classification, visualization, and frequency analysis of translation errors in Sinhala-English Machine Translation systems. The tool identifies frequent translation errors to enhance system performance.

Stack : Python, NLTK, Jupyter Notebook.

### **Improving Sinhala - English Machine Translation quality using Automatic Post Editing (APE) Techniques (2020)**

Research on current APE techniques and implementation of a suitable APE system to improve Sinhala - English Machine Translation quality

Stack : OpenNMT-py, OpenNMT-tf, Pytorch, LSTM, Transformer-Based Models.

### **Customer recommender solution for banking domain (2018)**

A customer recommender solution with a useful feature that enables the bank to describe customer behaviors and generate insight through them was designed and implemented using the customer transaction and demographic data available in a given bank database.

Stack : Python, PySpark, XGBoost, AWS virtual servers.

## PROPOSAL WRITING EXPERIENCE

---

### **NSF Growing Convergence Research (GCR) Proposal**

Submitted 2026 January

National Science Foundation

Role: Contributed to establishing the technical foundation and intellectual framework of the proposal; provided preliminary results (BLADE)

### **NIH R01 Proposal on Scalable Pre-consultation Frameworks in Healthcare**

Submitted 2026 March

National Institutes of Health

Role: Contributed to proposal writing (scalability extension of existing framework (BLADE) in patient consultation setting)

## TALKS AND PRESENTATIONS

---

### **Oral Presentations**

- [From Disagreement to Understanding: The Case for Ambiguity Detection in NLI](#), The fourth Workshop on Perspectivist Approaches to NLP (NLPerspectives) - EMNLP 2025, Suzhou, China (Virtual), November 2025

### **Poster Presentations**

- [AMREx: AMR for Explainable Fact Verification](#), The seventh Fact Extraction and VERification Workshop (FEVER) - EMNLP 2024, Miami, Florida, USA, November 2024

## ACADEMIC SERVICES

---

### **Conference Reviewing**

- **ACL Rolling Review (ARR):** Area Chair, 2025 May Cycle

### **Mentoring**

- **FINS Talent Pipeline:** Graduate Mentor, 2025
- **NLP&C Laboratory:** Graduate Mentor, 2023 - Present

## OPEN-SOURCE CONTRIBUTIONS

---

### **Project Aya by Cohere for AI**

Contributed to help developing instruction-style datasets for Sinhala, a low-resourced language, as part of their initiative to create a massively multilingual LLM supporting under-represented languages.

### **Fine-tuning BERT and DistilBERT models in HuggingFace**

Fine-tuned two BERT-base-uncased models for analyzing political discourse and Abstract Meaning Representation (AMR) sequence processing

## SKILLSET

---

Programming Languages: Python | R | C++  
Softwares : Matlab | Weka  
Machine Learning Frameworks: Pytorch | HuggingFace | Opennmt-py  
NLP Tools: Chroma | Faiss | Spacy | NLTK | XGBoost | Transformers  
Technical Skills : LLM-based Evaluation, Data Visualization, Preprocessing, Analysis and Modeling, R&D, Conversational AI  
Soft Skills : Collaborative Work | Problem-Solving | Self-Starter | Research Communication  
Other : Git | CSS | Flask | LaTeX | AMR graphs

## AWARDS & ACHIEVEMENTS

---

2023: CISE Scholarship Award, Gartner Group Graduate Fellowship for excellence in publication record, University of Florida  
2015-2019: Dean's Honour List for outstanding academic achievements in Semester 1, 2, 6, 7 and 8.  
2015-2019: Mahapola merit scholarship for outstanding performance in GCE Advanced Level by the government of Sri Lanka.  
2018: India's first Bank-led Datathon organized by Yes Bank, India - 1st runner up and Best Student team award  
2018: Hackstat 2018, a Datathon organized by the Stat Circle of University of Colombo - 1st Place  
2014: Sri Lankan Physics Olympiad Competition - Merit Award

## REFERENCES

---

Available upon request