Curriculum Vitae

Personal Information

- Full name: Gabriel Emerson Iturra Bocaz

- Nationality: Chile

- E-mail: gabrieliturrab at ug.uchile.cl / giturra at dcc.uchile.cl

- Public Profiles: personal website, IMFD, Google Scholar, LinkedIn, GitHub, ResearchGate, Semantic Scholar, Orcid, Twitter.

- Areas of Interest: Natural Language Processing, Large Language Models, Data Stream Mining, Machine Learning, Artificial Intelligence

Short Bio

Large Language Models (LLMs) are significant deep learning models trained on massive corpora of text. These models have revolutionized the field of Natural Language Processing (NLP) with their ability to accurately solve various NLP tasks, such as sentiment analysis, text classification, and text summarization. These breakthroughs have prompted questions about whether these models can learn new tasks without retraining with new data, or if the vast amount of data they were trained on has unveiled a new characteristic in AI models: the ability to reason. However, despite their revolutionary breakthroughs, LLMs lack the ability to solve planning tasks—a capability highly desired by Autonomous Agents. This gap is why I am deeply interested in approaches that enhance the planning abilities of autonomous agents using the semantic knowledge that LLMs can provide. During my master's degree, I developed RiverText, a framework for training and evaluating word embeddings from text data streams. This framework utilizes multiple methods and algorithms to incrementally represent word embeddings based on the distributional hypothesis. My work received international recognition, being accepted at the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '23) in Taipei, Taiwan. Currently, I serve as a part-time professor in the Master of Data Science program at the University of Chile, where we delve into the most crucial aspects of Machine Learning and Data Mining. I am also actively pursuing PhD opportunities to delve deeper into the potential of Large Language Models. My goal is to improve their prompting engineering capabilities, expand their functionalities with Reinforcement Learning methods, and leverage their capabilities across various disciplines.

Education

- MSc. Computer Science, University of Chile (2023)

- Computer Science Engineer, University of Chile (2021)

- BSc. Computer Science, University of Chile (2020)
- BSc. in Industrial Engineering, University of Chile (2019)

Experience

- August 2023 - Present: Part-Time Professor on Master of Data Science, University of Chile.
- Jul 2021 - Jun 2023 - Present: Research Assistant, FONDECYT 11200290, Chile. Project: Tracking Social Public Opinion: A Stream-Mining Based Approach.
- Jul 2021 - Jun 2023: Researcher Intern, Millennium Institute for Foundational Research on Data, Chile.
- Jul 2021 - Jun 2023: Researcher Intern, National Center for Artificial Intelligence, Chile.
- March 2018 - September 2020: Director of I. T., Beauchef Networking.
- March 2020 - July 2020: Web Developer, UGU S. A.
- January 2019 - March 2019: Web Developer, Memory and Neuropsychiatry Clinic, Faculty of Medicine, University of Chile.
- January 2018 - February 2018: Intership, Wizz, Mobile Developer.
- December 2016 - January 2017: Intership, Web Intelligence Centre, Software Developer.

Teaching

- [MDS7202] (Lecturer) Data Science Laboratory (spring 2023), Postgraduate School, Universidad de Chile.
- [CC66S] (TA) Machine Learning (Spring 2023), Continual Education Program, Department of Computer Science, Universidad de Chile.
- [CC66Q] (TA) Natural Language Processing (spring 2023), Department of Computer Science, Universidad de Chile.
- [DEDCCPYAC02] (TA) Data Management and Data Mining in Python (spring 2023), Continual Education Program, Department of Computer Science, Universidad de Chile.
- [CC5402] (TA) Software Project (spring 2023), Department of Computer Science, Universidad de Chile.
- [CC6105] (TA) Statistical Thinking (spring 2022), Department of Computer Science, Universidad de Chile.

- [CC6204] (TA) Deep Learning (spring 2022), Department of Computer Science, Universidad de Chile.

- [CC5205] (TA) Data Mining (fall 2022, fall 2023), Department of Computer Science, Universidad de Chile.

- [CC6205] (TA) Natural Language Processing (fall 2022, fall 2023), Department of Computer Science, Universidad de Chile.

- [CC7910] (TA) Research in Computer Science (Methods, Techniques, Perspectives) (spring 2021, fall 2022, spring 2022, fall 2023, spring 2023), Department of Computer Science, Universidad de Chile.

- [IN6534] (TA) Introduction to Deep Learning (spring 2021), Department of Industrial Engineering, Universidad de Chile.

- [CC4102] (TA) Algorithm Design and Analysis (spring 2020, fall 2022, spring 2022, fall 2023), Department of Computer Science, Universidad de Chile.

- [CC4101] (TA) Programming Languages (spring 2020), Department of Computer Science, Universidad de Chile.

- [CC5401] (TA) Software Engineering II (fall 2020), Department of Computer Science, Universidad de Chile.

- [CC3201] (TA) Databases (fall 2019, spring 2019), Department of Computer Science, Universidad de Chile.

- [IQ5412] (TA) Process Modeling and Simulation (spring 2018), Department of Chemical Engineering and Biotechnology, Universidad de Chile.

- [CC5002] (TA) Web Application Development (spring 2017, fall 2018, spring 2018, fall 2019, spring 2019, fall 2020,spring 2020, fall 2021, spring 2021), Department of Computer Science, Universidad de Chile.

- [MA2601] (TA) Ordinary differential equations (fall 2016), Department of Mathematical Engineering, Universidad de Chile.

Short Courses

- LSSDS, La Serena School Data Science, 2-13 August, 2021, La Serena (Certificate of Attendance).

Publications

1. G. Iturra-Bocaz and F. Bravo-Marquez RiverText: A Python Library for Training and Evaluating Incremental Word Embeddings from Text Data Stream. In Proceedings of the 46th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2023),

Taipei, Taiwan. Association for Computing Machinery. Pages 3027–3036. DOI:10.1145/3539618.3591908 (pdf)(poster).

2. F. Vera, V. D. Cortés, G. Iturra, J. D. Velásquez, P. Maldonado and A. Couve, Akori: A Tool Based in Eye-Tracking Techniques for Analyzing Web User Behaviour on a Web Site 2017 IEEE International Conference on Data Mining Workshops (ICDMW), 2017, pp. 635-640, doi: 10.1109/ICDMW.2017.90 (pdf).

Thesis

1. Master's Thesis (first five pages in Spanish): RiverText A framework for training and evaluating Incremental Word Embeddings from Text Data Streams (pdf).

Projects

1. RiverText: A Python Library for Training and Evaluating Incremental Word Embeddings from Text Data Streams (as main developer).

2. MDS7202 Official repository from the course Data Science Laboratory at the Data Science Master Program at the University of Chile (as lecturer).

3. CC66Q Official repository for the Natural Language Processing module in the Artificial Intelligence Diploma Program at the University of Chile (as Teaching Assistant).

Awards

- [2023] SIGIR Student Travel Grant to attend the 2023 SIGIR International Conference on Information Retrieval in Taipei, Taiwan.

- [2016] Outstading Student in 2016.

Talks and Workshops

- [September 2023] Poster presentation RiverText: A Framework for Training and Evaluating Incremental Word Embeddings from Text Data Streams, The first Postgraduate Symposium at the Faculty of Math and Physics Sciences from the University of Chile, Santiago, Chile (Certificate of Attendance).

- [July 2023] Paper Presentation entitled: RiverText: A Python Library for Training and Evaluating Incremental Word Embeddings from Text Data Streams, The 46th International ACM SIGIR Conference on Research and Development in Information Retrieval, Taipei, Taiwan (Certificate of Attendance).

- [July 2023] Workshop: Web Scrapping, The Summer Institutes in Computational Social Science, Santiago, Chile.

- [July 2023] Workshop: Text Analysis II, The Summer Institutes in Computational Social Science, Santiago, Chile.

- [April 2023] Department of Computer Science, Universidad de Chile. Thesis Talk, entitled RiverText: A Framework for Training and Evaluating Incremental Word Embeddings from Text Data Streams (slides)(video in Spanish).

Media Outreach

- Revista Bits N° 25: Inteligencia Artificial Generativa, Estudiantes, página 63.

Technical Skills

- Programming Languages: Python, Java, C/C++, R, PHP, JavaScript.

- Especial Frameworks: PyTorch, TensorFlow, Django, Springs Boots, React and React Native.

- Databases: PostgreSQL, MySQL and MongoDB.

Languages

- Spanish: Native.

- English: Advance (B2), Academic IETLS 6.5/9.0.

References

- Felipe Bravo-Marquez: Assistant Professor, University of Chile. Contact: fbravo at dcc.uchile.cl

- Claudio Gutierrez: Full Professor, University of Chile. Contact: cgutierr at dcc.uchile.cl

Gabriel Iturra-Bocaz

Santiago, Chile September 9, 2023