

José Cañete López

MSc in Computer Science student from Universidad de Chile.

Interested in applied research of ML / NLP, especially Transfer Learning and Model Compression in order to bring the advances of NLP to underrepresented languages and countries.

Education

- 2021 - 2023** **MSc in Computer Science;** Universidad de Chile (Santiago, Chile)
- 2014 - 2021** **BSc, Computer Science Engineering;** Universidad de Chile (Santiago, Chile)

Work Experience

- March 2022 - Now (Part-time)** **Machine Learning Engineer at National Center for Artificial Intelligence (CE-NIA)** (Santiago, Chile).
In charge of develop and make it into production AI products for different businesses or research projects.
- Jan. 2021 - Nov. 2021 (Full-time)** **Lead Data Scientist at Adereso** (Santiago, Chile).
Technical lead of the Adereso Minds team, in which I was in charge of a Data Scientist and a Full Stack Engineer. At Minds we developed projects related to data and artificial intelligence, from conception and prototype to production, using agile development methodologies and cloud technologies.
- Jan. 2019 - Dec. 2020 (Part-time)** **Data Scientist at Adereso** (Santiago, Chile).
Developed, experimented and deployed into production different models of Machine Learning applied to text (NLP). In particular trained models like BERT and its variants (ALBERT, S-BERT, among others) for understanding Spanish informal text messages coming from social networks. Technologies: Python, PyTorch, Pandas, Tornado, BigQuery, RabbitMQ, AWS, GCP.

Research Experience

- 2022** **Two-stage Conditional Chest X-ray Radiology Report Generation** in MedNeurIPS 2022, New Orleans, USA.
Pablo Messina, **José Cañete**, Denis Parra, Alvaro Soto, Cecilia Besa, Jocelyn Dunstan. [Paper](#), [Repo](#).
- 2022** **ALBETO and DistilBETO: Lightweight Spanish Language Models** in LREC 2022, Marseille, France.
José Cañete, Sebastián Donoso, Felipe Bravo-Marquez, Andrés Carvallo, Vladimir Araujo. [Paper](#), [Repo](#).

- 2022** **Evaluation Benchmarks for Spanish Sentence Representations** in LREC 2022, Marseille, France.
- Vladimir Araujo, Andrés Carvallo, Souvik Kundu, **José Cañete**, Marcelo Mendoza, Robert E. Mercer, Felipe Bravo-Marquez, Marie-Francine Moens, Alvaro Soto. **Paper**, **Repo**.
- 2020** **Spanish Pre-trained BERT Model and Evaluation Data** in PML4DC @ ICLR 2020, Addis Ababa, Ethiopia.
- José Cañete**, Gabriel Chaperon, Rodrigo Fuentes, Jou-Hui Ho, Hojin Kang and Jorge Perez. **Paper**, **Repo**.

Additional Experience and Awards

- Conferences and Summer Schools**
- Language Resources and Evaluation Conference** - Marseille, France, 2022.
 - IV Workshop Millenium Institute for Foundational Research on Data (IMFD)** - Santa Cruz, Chile, 2022.
 - International Conference on Learning Representations** - Virtual, 2020.
 - Khipu AI - Latin American Meeting in Artificial Intelligence** - Montevideo, Uruguay, 2019.
 - II Workshop Millenium Institute for Foundational Research on Data (IMFD)** - Viña del Mar, Chile, 2019.
 - Brazilian ICPC Summer School** - Campinas, Brasil, 2017.
 - Winter Competitive Programming Camp Chile (ICPCCL)** - Chile, 2015 and 2016.
- Awards**
- Outstanding Student in Research**, DCC UChile, 2019-2020.
 - 1st Place**, Winter Competitive Programming Camp of Chile (ICPCCL) 2017.
 - 3rd Place**, Chilean Programming Contest 2017.
 - 3rd Place**, ACM ICPC Chile 2017.
 - 9th Place and First to Solve**, ACM ICPC Chile 2016.

josecannete@gmail.com • +56 9 8438 7585 • github.com/josecannete
Antuhuenu 7216 - Santiago, Chile