

# José Cañete López

---

## Estudiante de Magíster en Ciencias de la Computación de la Universidad de Chile.

Interesado en la investigación aplicada en ML y NLP, especialmente en Transfer Learning y Model Compression, con el fin de traer los avances en NLP a lenguajes y países subrepresentados.

---

## Educación

- |             |  |
|-------------|--|
| 2021 - 2023 | <b>Magíster en Ciencias, mención Computación;</b> Universidad de Chile (Santiago, Chile) |
| 2014 - 2021 | <b>Ingeniería Civil en Computación;</b> Universidad de Chile (Santiago, Chile)           |

## Experiencia Laboral

- |  |  |
|--|--|
| Marzo 2022<br>- Ahora (Part-time)      | <b>Machine Learning Engineer at National Center for Artificial Intelligence (CE-NIA)</b> (Santiago, Chile).<br><br>Encargado de desarrollar y llevar a producción productos de IA para clientes o proyectos de investigación.  |
| Enero 2021<br>- Nov. 2021 (Full-time)  | <b>Lead Data Scientist en Adereso</b> (Santiago, Chile).<br><br>Líder técnico del equipo Adereso Minds, en el cual estuve a cargo de un Data Scientist y un Full Stack Engineer. En Minds desarrollabamos proyectos relacionados a datos e inteligencia artificial, desde su concepción y prototipo hasta su puesta en producción, haciendo uso de metodologías ágiles de desarrollo y tecnologías cloud. En particular: <ul style="list-style-type: none"><li>• Se entrenó modelos del estado del arte de NLP como por ejemplo BERT, RoBERTa, ALBERT, entre otros.</li><li>• Se usaron técnicas como Knowledge Distillation para hacer estos modelos más eficientes y más rápidos a la hora de ponerlos en producción.</li><li>• Se hizo uso de herramientas de manejo y procesamiento de grandes volúmenes de datos como Apache Beam (tanto para procesamiento en batch como en streaming), BigQuery y PubSub.</li><li>• Se usó tecnologías serverless como Cloud Functions y Cloud Run para la puesta en producción de nuestras APIs que servían los modelos de ML.</li><li>• Se usó extensivamente Docker y herramientas como Drone y Cloud Build para el CI/CD de nuestros proyectos.</li></ul> |
| Enero 2019<br>- Sept. 2020 (Part-time) | <b>Data Scientist en Adereso</b> (Santiago, Chile).<br><br>Desarrollo, experimentación y puesta en producción de diversos modelos de Machine Learning aplicados a texto (NLP). En particular se implementó, entrenó e iteró con modelos de tipo BERT y sus variantes (ALBERT, S-BERT, entre otros)   |

aplicados a mensajes en español informal, proveniente de redes sociales. Tecnologías: Python, PyTorch, Tornado, BigQuery.

## Experiencia en Investigación

- 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](#).

## Experiencia Adicional y Premios

- Conferencias y Escuelas de Verano**
- Language Resources and Evaluation Conference** - Marseille, France, 2022.
  - IV Workshop Instituto Milenio Fundamentos de los Datos (IMFD)** - Santa Cruz, Chile, 2022.
  - International Conference on Learning Representations (ICLR)** - Virtual, 2020.
  - Khipu AI - Latin American Meeting in Artificial Intelligence** - Montevideo, Uruguay, 2019.
  - II Workshop Instituto Milenio Fundamentos de los Datos (IMFD)** - Viña del Mar, Chile, 2019.
  - Brazilian ICPC Summer School** - Campinas, Brasil, 2017.
  - Invernal Campamento de Programación Competitiva en Chile (ICPCCL)** - Chile, 2015, 2016 y 2017.
- Premios**
- Estudiante Destacado en Investigación Internacional**, DCC UChile, 2019-2020.
  - Primer Lugar**, Invernal Campamento de Programación Competitiva en Chile (ICPCCL) 2017.
  - Tercer Lugar**, Torneo Chileno de Programación (TCP) 2017.

**Tercer Lugar**, ACM ICPC Chile 2017.

**9º Lugar y First to Solve**, ACM ICPC Chile 2016.

---

[josecannete@gmail.com](mailto:josecannete@gmail.com) • +56 9 8438 7585 • [GitHub](#) • [LinkedIn](#)

Antuhuenu 7216 - Santiago, Chile