Jean-Baptiste LAMARE
Research Scientist | Deep Learning and NLP

in linkedin.com/in/jlamare **O** github.com/jblamare **⚲** Google Scholar **↗** Personal Website
▯ 412-626-8707 **@** jblamare@gmail.com **◈** San Francisco, CA **i** French citizen, H-1B visa

🎓 EDUCATION

CARNEGIE MELLON UNIVERSITY (CMU), SCHOOL OF COMPUTER SCIENCE - PITTSBURGH, PA   2017 - 2019
M.S., Language Technologies - Focus : Natural Language, Computer Vision, Audio Processing

ECOLE POLYTECHNIQUE, DEPARTMENT OF COMPUTER SCIENCE - PALAISEAU, FRANCE   2012 - 2017
M.S., Data Science - Focus : Machine Learning, Databases, Text Mining and NLP
B.S., Mathematics and Computer Science

💼 PROFESSIONAL EXPERIENCE

ENLITIC
San Francisco, CA

SENIOR RESEARCH SCIENTIST (April 2021 - Present)   2019 - PRESENT
RESEARCH SCIENTIST (August 2019 - April 2021)

- Parsed and analyzed millions of **EHR and DICOM data points** to extract insightful features in order to streamline the work of radiologists.
- Developed self-supervised, supervised and unsupervised NLP models for **multiclass/multilabel classification, NER, and entity linking**, using various architectures (transformers, BERT, RNNs...) to detect over 70 abnormalities in clinical reports, extract entities and match them to our ontology.
- Implemented several NLP models and pipelines for **free text report structuring**, allowing the identification of sections relevant to each target feature.
- Trained supervised multilabel **computer vision models**, as well as **multimodal models** at the intersection of CV and NLP, to detect over 20 abnormalities in radiological images, specifically X-rays.
- Designed end-to-end labeling pipelines for reports and images, including **interface design, case selection, quality assurance** that would be used by over 15 experts to label more than 100,000 studies.
- Developed and maintained several **shared repositories** and **python packages** for deep learning model training (Keras-like), model deployment, as well as NLP pipeline creation (spaCy-like).
- Led a **successful blind test** demonstrating model outputs on independently managed datasets for prospective clients.

CMU
Pittsburgh, PA

GRADUATE RESEARCH ASSISTANT   2017 - 2019

- Collected and labeled a **dataset of tweets** during public safety events.
- Developed a **Deep Learning useful tweets extraction** system for public safety events to identify tweets that contain information helping the authorities respond to such crises.
- Designed and implemented a **web interface** for users to visualize extracted tweets in real-time and provide feedback on which tweets are actually useful, feeding an active learning paradigm.

📖 PUBLICATIONS

EMNLP 2020
Clinical NLP WS

ON THE DIMINISHING RETURN OF LABELING CLINICAL REPORTS   NOVEMBER 2020
Jean-Baptiste Lamare, Tobi Olatunji, Li Yao

AVSS 2018
T4S WS

ACCIDENT FORECASTING IN CCTV TRAFFIC CAMERA VIDEOS   NOVEMBER 2018
Ankit Shah, Jean-Baptiste Lamare, Tuan Nguyen Anh, Alexander Hauptmann

ICMR 2018

MULTIMODAL FILTERING OF SOCIAL MEDIA FOR TEMPORAL MONITORING AND EVENT ANALYSIS   JUNE 2018
Po-Yao Huang, Junwei Liang, Jean-Baptiste Lamare, Alexander Hauptman

💻 SELECTED PROJECTS

EMBODIED QUESTION ANSWERING (EQA)   AUGUST 2018 - DECEMBER 2018

- Improved the Visual Question Answering (VQA) and navigation metrics on the EQA task.
- Worked on deep learning applied to VQA and navigation in a 3D environment in Pytorch.

A NEURAL MUSIC CRITIC   JANUARY 2018 - MAY 2018

- Created a model that automatically writes a review from the audio file of a song.
- Worked on Music Processing and Conditioned Text Generation with CNNs and LSTMs in Pytorch.

☰ SKILLS

Programming Languages	Python, Java, C++, R, SQL, HTML, Javascript, CSS
Machine Learning frameworks	PyTorch, Tensorflow, spaCy, Stanza, Scikit-learn, Pandas, Numpy
Other Tools/Technology	Git, AWS, Elasticsearch, Mechanical Turk, Raspberry
Languages	English (fluent), French (native), German (intermediate), Spanish (beginner)

🚀 MISCELLANEOUS

Hobbies   Competitive tennis and skiing, guitar, piano, chess, baking, traveling.