Muammar W. El Khatib Rodriguez, Ph.D.

Work Address
Bristol Myers-Squibb.
Informatics and Predictive Sciences.
250 Water St.
Cambridge, MA 02141 United States.
Tel.: +1 (401) 450-1980.
muammar.elkhatib@bms.com

Personal Contact
Somerville, MA 02145 United States.
Cel.: +1 (401) 450 1980.
muammarelkhatib@gmail.com
`https://github.com/muammar`
`https://stackoverflow.com/users/1995261/muammar`
`https://www.linkedin.com/in/muammarelkhatib`
`https://muammar.me`

CORE SKILLS

Python | Pytorch | Dask | RAPIDS | Keras | Scikit-learn | JupyterLab | Pandas | MongoDB | BASH Scripting | Debian Packaging | Linux | macOS | Git | Research | Deep learning | Kernel methods | Software development | Physics | Quantum Chemistry | Matplotlib | Plotly | Seaborn | Docker | Data Analysis | Graph Neural Networks | Uncertainty Quantification | Generative Models | Cheminformatics | Supervisory Skills | Management

WORK EXPERIENCE

Bristol Myers-Squibb — January 2021 to March 2024.
Informatics and Predictive Sciences organization. — (Principal Machine Learning Scientist)
250 Water St., Cambridge, MA 02141, USA.

Bristol Myers-Squibb — November 2020 to January 2021.
200 Cambridgepark Dr., Cambridge, MA 02140, USA.
Informatics and Predictive Sciences organization. — (Senior Machine Learning Scientist)

A collaborative, creative, and interdisciplinary role in applied machine-learning research in partnership between Informatics & Predictive Sciences (IPS) and Chemistry organizations. Responsibilities include:

- Report to the lead for Predictive Sciences of the IPS organization; scenarios involve a range of datasets and learning objectives, including drug discovery, multi-task and multi-modal modeling, uncertainty quantification, and predictive models for chemical and biological datasets.
- Formulation and implementation of predictive modeling and machine learning solutions for optimizing chemical structures, drug-target interactions, and physicochemical properties.
- Application of cutting-edge machine learning (deep learning) approaches to physicochemical property predictions and molecular interaction challenges.
- Work alongside experts in similar applications of machine learning in the biotechnology domain, including:

    1. Collaboration to develop human-in-the-loop systems to capture and operationalize machine learning datasets and algorithms used by BMS scientists.
    2. Application of supervised and semi-supervised deep learning methods to derive robust, generalizable, and reusable representations for chemical and biological assay data.
    3. Design of multi-task, multi-modal, and generative neural network learning approaches to tackle real-world drug discovery optimization problems.

- Pursue leading research in applied scientific machine learning that demonstrates the value of predictive methods to accelerate and optimize drug development.

- Derive and apply predictive approaches in collaboration with BMS colleagues in the Informatics Predictive Sciences and Chemistry departments.
- Apply rigorous internal standards for applied machine learning practice, including evaluation of methods, approaches, and solutions.
- Present strategies, techniques, results, and conclusions to BMS colleagues and external audiences.
- Enable strategic collaborations with academic and commercial collaborators to benefit therapeutic programs. I am the BMS representative before the "Machine Learning for Pharmaceutical Discovery and Synthesis Consortium" collaboration with the Massachusetts Institute of Technology. For more information, see `https://mlpds.mit.edu/`.
- Implement and apply uncertainty quantification methods of predictive models to steer scientific discovery.
- Lead and supervise (in)direct reports to develop their careers and benefit therapeutic programs at BMS.

Lawrence Berkeley National Laboratory November 2018 to September 2020.
Computational Research Division. (Postdoctoral Scholar)
1 Cyclotron Road, MS 50F-1650 Berkeley, CA 94720, USA.

- Developed a Python library to ease the deployment of machine learning models for chemistry and materials sciences. `https://github.com/muammar/ml4chem`. This package is helping us advance our research faster because we can consistently implement new methods.
- Developed a neural network model that can learn how to predict retention times from chromatography data. Experimentalists can use these models to get insights into the substances they study without explicitly running the experiments.
- Applied autoencoders to extract features and systematically studied their topology to understand their effect on the predictive power of models used for material sciences.
- Generated data sets using web-scraping and diversified their variance with active learning techniques.
- Worked on a project for Scaling Interactive Science for Data-Intensive Discovery for the Linac Coherent Light Source (LCLS) at SLAC National Accelerator Laboratory.
- Led and designed research projects executed by summer interns.
- Wrote scientific publications to show our results to the community.

Brown University October 2016 to November 2018.
School of Engineering. (Postdoctoral Research Associate)
Box D. Brown University. 184 Hope Street. Providence, RI 02912 USA.

- Worked on the acceleration of electronic structure calculations using machine learning models to decrease orders of magnitude of the computational time needed by the simulations.
- Was actively involved in the development of the Atomistic Machine-learning Package (Amp) created and maintained by the Catalyst Design Laboratory at Brown University `https://bitbucket.org/andrewpeterson/amp`.
- Implemented kernel ridge regression within an atom-centered mode in their machine-learning package.
- Participated in the design of scientific projects and supervision of students during their research in our laboratory.
- Presented scientific results at international conferences.

Debian Project 2006 to present.
`http://www.debian.org`

- Debian Developer (muammar@debian.org). For a list of packages maintained by me, please visit: `http://qa.debian.org/developer.php?login=muammar`

EDUCATION

Ph.D. in Theoretical Chemical Physics July 2012 to July 2015.
Université Paul Sabatier III, Toulouse, France.

Master on Theoretical Chemistry and Computational Modelling September 2010 to July 2012.
Université Paul Sabatier III, Toulouse, France.

LIST OF PUBLICATIONS

Below are my most recent scientific publications. For a complete list, please see `https://scholar.google.com/citations?user=WBXXCScAAAAJ&hl=en`.

1. X.Chen , **M. El Khatib**, P. Lindgren , A. Willard , A. Medford, A. Peterson, "Atomistic learning in the electronically grand-canonical ensemble," npj Comput Mater 9, 73 (2023). DOI: 10.1038/s41524-023-01007-6
2. **M. El Khatib**, "ML4Chem: A Machine Learning Package for Chemistry and Materials Science", arXiv:2003.13388 (2020).
3. M.D. Hanwell, WA. De Jong, J. Hachmann, C. Harris, A. Genova, **M. El Khatib**, P. Avery, "Open Chemistry, JupyterLab, REST, and Quantum Chemistry," International Journal of Quantum Chemistry, 121(1), e26472 (2020). DOI: 10.1002/qua.26472

LANGUAGES & INTERESTS

Spanish, Native language | **English**, Full professional proficiency | **French**, Full professional proficiency.

Classical guitar | Free software community | Physics | Electronic structure theory | Technologies.