David Alvarez-Melis

Assistant Professor of Computer Science
Harvard John A. Paulson School of Engineering and Applied Sciences

150 Western Av., #2.332
Allston, MA, 02134
✉ *dam[at]seas[dot]harvard[dot]edu*

Research Interests

Themes — Geometry in machine learning, transfer + multi-domain learning, interpretability.

Methods — Optimal transport, convex/submodular optimization, differential equations.

Applications — Natural language processing, medical imaging, biochemistry, scientific discovery.

Research and Work Experience

2023 – **Assistant Professor**, *Harvard University*, Allston, MA, USA
- Primary affiliation: John A. Paulson School of Engineering and Applied Sciences (SEAS).
- Secondary Harvard affiliations: Kempner Institute, Harvard Data Science Initiative (HDSI), Center for Research on Computation and Society (CRCS).

2021 – **Senior Researcher**, *Microsoft Research*, Cambridge, MA, USA

2019 – 2021 **Postdoctoral Researcher**, *Microsoft Research*, Cambridge, MA, USA
- Topics: optimal transport for meta-learning, debiasing and adaptation

2014 – 2019 **Research Assistant**, *MIT CSAIL*, Cambridge, MA, USA
- Supervisor: Tommi Jaakkola.
- Recent Projects: structured optimal transport, robustly interpretable machine learning.

05 – 08/2018 **Research Intern**, *Microsoft Research*, New York, NY, USA
- Mentors: Hanna Wallach, Jenn Wortman Vaughan, Hal Daumé III.
- Project: Robust and human-like interpretability for machine learning.

05 – 08/2016 **Research Intern**, *Microsoft Research*, Redmond, WA, USA
- Mentors: Scott Yih, Ming-Wei Chang, Kristina Toutanova, Chris Meek.
- Project: Multi-hop relation prediction for knowledge base question answering.

2013 – 2014 **Supplemental Researcher**, *IBM Research*, TJ Watson Center, NY, USA
- Mentors: Michael Picheny & Ken Church (speech recognition group).
- Data mining, statistical modeling and machine learning for speech recognition data.

2009 – 2010 **Statistical Analyst**, *LasQuinceLetras Solutions*, Mexico City, Mexico
- Designed and carried out statistical learning methods on large survey datasets.

Education

2014 – 2019 **Massachusetts Institute of Technology**, Ph.D in Computer Science
- Area: Machine Learning, minor in Mathematical Optimization.
- Thesis: *Optimal Transport in Structured Domains: Algorithms and Applications*
- Committee: Tommi Jaakkola (advisor), Stefanie Jegelka, Justin Solomon.

2011 – 2013 **Courant Institute, New York University**, M.S. in Mathematics
- Thesis: *The Matrix Multiplicative Weights Algorithm for Domain Adaptation.*
- Advisor: Mehryar Mohri.

2006 – 2011 **Instituto Tecnologico Autonomo de Mexico**, B.S. in Applied Mathematics
- Thesis: *The Lax-Milgram Theorem, Generalizations and Applications.*
- Advisor: Carlos Bosch Giral.
- Mención Honorífica (*summa cum laude*), top 1% of class, valedictorian.

Selected Publications

[S1] **D. Alvarez-Melis** and N. Fusi. "Dataset Dynamics via Gradient Flows in Probability Space". In: *Proc. 38th International Conference on Machine Learning*. Vol. 139. 2021.

[S2] **D. Alvarez-Melis**, H. Kaur, H. Daumé III, H. Wallach, and J. W. Vaughan. "From Human Explanation to Model Interpretability: A Framework Based on Weight of Evidence". In: *Proc. Ninth AAAI Conference on Human Computation and Crowdsourcing*. Vol. 9. 2021.

[S3] **D. Alvarez-Melis** and N. Fusi. "Geometric Dataset Distances via Optimal Transport". In: *Adv. Neural Information Processing Systems*. Vol. 33. 2020.

[S4] **D. Alvarez-Melis**, T. Jaakkola, and S. Jegelka. "Structured Optimal Transport". In: *Proc. Twenty-First International Conference on Artificial Intelligence and Statistics*. Vol. 84. 2018.

[S5] **D. Alvarez-Melis** and T. S. Jaakkola. "Towards Robust Interpretability with Self-Explaining Neural Networks". In: *Adv. Neural Information Processing Systems*. Vol. 31. 2018.

Grants, Fellowships and Awards

2024 **Aramont Fellowship for Emerging Science Research**, Aramont Fellowship Fund
2023 **Dean's Competitive Fund for Promising Scholarship**, FAS, Harvard University
2023 **Top Reviewer Award**, *AISTATS 2023*
2021 **Top Reviewer Award**, *ICLR 2021*
2020 **Outstanding Reviewer Award**, *ICML 2020*
2019 **Best Reviewer Award**, *NeurIPS 2019*
2018 **Facebook Fellowship Finalist**, (30/800 applicants)
2018 **Best Reviewer Award**, *NeurIPS 2018*
2018 **Hewlett Packard Graduate Fellowship**, One-term PhD award
2018 **AI2 Key Scientific Challenges program award**, $10K unrestricted award
2011, 2014 **Fellowship for graduate studies abroad**, *CONACYT*
2012 **Alumni Research Prize**, *ITAM*, Category: Undergraduate Thesis
2011 **Sotero Prieto Prize, Second Place**, *Mexican Mathematical Society*
2006 – 2009 **Academic Excellence Scholarship**, *ITAM*, For undergraduate studies

Press and Outreach

2020 **Microsoft Research Blog**, "Measuring dataset similarity using optimal transport"
2019 **ZDNet**, "IBM offers explainable AI toolkit, but it's open to interpretation"
2018 **MIT News**, "Model paves way for faster, more efficient translations of more languages"
2018 **VentureBeat**, "MIT CSAIL is using unsupervised learning for language translations"
2017 **MIT News**, "How Neural Networks think"

Professional Activities and Service

Reviewer ACL, IJCNLP, UAI, NeurIPS, ICML, ICLR, AISTATS, LXAI, OTML, PLoS ONE, JAIR, TACL, JMLR, TMLR, IMAIAI, TPAMI, AIJ, *Nature Human Behavior*, SIMODS.

AC/Meta-R. ICML 2022, ACML 2022, ACML 2023, NeurIPS 2023, ICLR 2024.

Organizer ICML 2022 (Associate Chair); Optimal Transport and Machine Learning Workshop at NeurIPS 2023; LatinX in AI Workshop at ICML 2023.

Organizer RIIAA 2018 (student-run AI conference in Mexico City), riiaa.org.

Organizer MLXMIT: Machine Learing across MIT (2019).

Other MIT EECS Graduate Admissions Committee (2017, 2019).

Other Orientation Co-Chair, MIT Graduate Student Council.

Full List of Publications

Most recent publications via Google Scholar.

Preprints and Tech Reports

[Pr1] **D. Alvarez-Melis**, N. Fusi, L. Mackey, and T. Wagner. "Budget-Constrained Bounds for Mini-Batch Estimation of Optimal Transport". In: (2022). arXiv: 2210.13630 [cs.LG].

[Pr2] **D. Alvarez-Melis** and T. Broderick. "A translation of "The characteristic function of a random phenomenon" by Bruno de Finetti". In: (2015). arXiv: 1512.01229 [math.ST].

Conference and Journal Publications

[C1] C.-Y. Chuang, S. Jegelka, and **D. Alvarez-Melis**. "InfoOT: Information Maximizing Optimal Transport". In: *Proc. 40th International Conference on Machine Learning*. Vol. 202. 2023.

[C2] K. Falahkheirkhah, A. Lu, **D. Alvarez-Melis**, and G. Huynh. "Domain adaptation using optimal transport for invariant learning using histopathology datasets". In: *Proceedings of The 6th International Conference on Medical Imaging with Deep Learning*. 2023.

[C3] J. Fan and **D. Alvarez-Melis**. "Generating Synthetic Datasets by Interpolating along Generalized Geodesics". In: *Proc. Thirty-Ninth Conference on Uncertainty in Artificial Intelligence*. Vol. 216. 2023.

[C4] **D. Alvarez-Melis**, V. Garg, and A. Kalai. "Are GANs overkill for NLP?" In: *Adv. Neural Information Processing Systems*. Vol. 35. 2022.

[C5] **D. Alvarez-Melis**, Y. Schiff, and Y. Mroueh. "Optimizing Functionals on the Space of Probabilities with Input Convex Neural Networks". In: *Transactions on Machine Learning Research* (2022).

[C6] A. Yeaton, R. G. Krishnan, R. Mieloszyk, **D. Alvarez-Melis**, and G. Huynh. "Hierarchical Optimal Transport for Comparing Histopathology Datasets". In: *Proceedings of The 5th International Conference on Medical Imaging with Deep Learning*. Vol. 172. 2022.

[C7] **D. Alvarez-Melis** and N. Fusi. "Dataset Dynamics via Gradient Flows in Probability Space". In: *Proc. 38th International Conference on Machine Learning*. Vol. 139. 2021.

[C8] **D. Alvarez-Melis**, H. Kaur, H. Daumé III, H. Wallach, and J. W. Vaughan. "From Human Explanation to Model Interpretability: A Framework Based on Weight of Evidence". In: *Proc. Ninth AAAI Conference on Human Computation and Crowdsourcing*. Vol. 9. 2021.

[C9] **D. Alvarez-Melis** and N. Fusi. "Geometric Dataset Distances via Optimal Transport". In: *Adv. Neural Information Processing Systems*. Vol. 33. 2020.

[C10] **D. Alvarez-Melis**, Y. Mroueh, and T. Jaakkola. "Unsupervised Hierarchy Matching with Optimal Transport over Hyperbolic Spaces". In: *Proc. Twenty Third International Conference on Artificial Intelligence and Statistics*. Vol. 108. 2020.

[C11] **D. Alvarez-Melis**, S. Jegelka, and T. S. Jaakkola. "Towards Optimal Transport with Global Invariances". In: *Proc. Twenty-Second International Conference on Artificial Intelligence and Statistics*. Vol. 89. 2019.

[C12] C. Bunne, **D. Alvarez-Melis**, A. Krause, and S. Jegelka. "Learning Generative Models across Incomparable Spaces". In: *Proc. 36th International Conference on Machine Learning*. Vol. 97. 2019.

[C13] G.-H. Lee, **D. Alvarez-Melis**, and T. S. Jaakkola. "Towards Robust, Locally Linear Deep Networks". In: *International Conference on Learning Representations*. 2019.

[C14] G.-H. Lee, W. Jin, **D. Alvarez-Melis**, and T. Jaakkola. "Functional Transparency for Structured Data: a Game-Theoretic Approach". In: *Proc. 36th International Conference on Machine Learning*. Vol. 97. 2019.

[C15] **D. Alvarez-Melis** and T. Jaakkola. "Gromov-Wasserstein alignment of word embedding spaces". In: *Proc. 2018 Conference on Empirical Methods in Natural Language Processing*. 2018.

[C16] **D. Alvarez-Melis**, T. Jaakkola, and S. Jegelka. "Structured Optimal Transport". In: *Proc. Twenty-First International Conference on Artificial Intelligence and Statistics*. Vol. 84. 2018.

[C17] **D. Alvarez-Melis** and T. S. Jaakkola. "Towards Robust Interpretability with Self-Explaining Neural Networks". In: *Adv. Neural Information Processing Systems*. Vol. 31. 2018.

[C18] **D. Alvarez-Melis** and T. S. Jaakkola. "A causal framework for explaining the predictions of black-box sequence-to-sequence models". In: *Proc. 2017 Conference on Empirical Methods in Natural Language Processing*. 2017.

[C19] **D. Alvarez-Melis** and T. S. Jaakkola. "Tree-structured decoding with doubly-recurrent neural networks". In: *Proc. International Conference on Learning Representations (ICLR)*. 2017.

[C20] **D. Alvarez-Melis** and M. Saveski. "Topic modeling in twitter: Aggregating tweets by conversations". In: *Proc. Tenth IEEE International Conference on Web and Social Media*. 2016.

Refereed Workshop Contributions

[W1] G. Giannone, N. Tenenholtz, J. Hall, N. Fusi, and **D. Alvarez-Melis**. "Enhancing Language Models for Technical Domains with Dynamic Token Injection". In: *Generative AI and Biology workshop (GenBio@NeurIPS2023)*. 2023.

[W2] A. Gupta, T. Moskovitz, **D. Alvarez-Melis**, and A. Pacchiano. "Undo Maps: A tool for Adapting Policies to Perceptual Distortions". In: *New Frontiers in Learning, Control, and Dynamical Systems Workshop at ICML*. 2023.

[W3] N. Hulkund, N. Fusi, J. W. Vaughan, and **D. Alvarez-Melis**. "Interpretable Distribution Shift Detection using Optimal Transport". In: *DataPerf Workshop at ICML*. 2022.

[W4] F. Lübeck, C. Bunne, G. Gut, J. S. del Castillo, L. Pelkmans, and **D. Alvarez-Melis**. "Neural Unbalanced Optimal Transport via Cycle-Consistent Semi-Couplings". In: *Learning Meaninful Representations of Life (LMRL) Workshop at NeurIPS*. 2022.

[W5] **D. Alvarez-Melis**, H. Daumé III, J. W. Vaughan, and H. Wallach. "Weight of Evidence as a Basis for Human-Oriented Explanations". In: *HCML: Workshop on Human-Centric Machine Learning at NeurIPS*. 2019.

[W6] H. James-Sorenson and **D. Alvarez-Melis**. "Probabilistic Bias Mitigation in Word Embeddings". In: *NeurIPS Workshop on Human-Centric Machine Learning*. 2019.

[W7] **D. Alvarez-Melis** and T. S. Jaakkola. "On the Robustness of Interpretability Methods". In: *Proc. 2018 ICML Workshop in Human Interpretability in Machine Learning*. 2018.

[W8] C. Bunne, **D. Alvarez-Melis**, S. Jegelka, and A. Krause. "Learning Generative Models Across Incomparable Spaces". In: *NeurIPS Workshop on Relational Representation Learning*. 2018.

[W9] G.-H. Lee, **D. Alvarez-Melis**, and T. S. Jaakkola. "Game-theoretic Interpretability for Temporal Modeling". In: *Fairness Accountability and Transparency in Machine Learning*. 2018.

[W10] **D. Alvarez-Melis** and J. Amores. "The Emotional GAN: Priming Adversarial Generation of Art with Emotion". In: *NeurIPS Workshop on Machine Learning for Creativity and Design*. 2017.

[W11] T. B. Hashimoto, **D. Alvarez-Melis**, and T. S. Jaakkola. "Word, graph and manifold embedding from Markov processes". In: *NIPS Workshop on Nonparametric Methods for Large Scale Representation Learning*. 2015.

Patents

[Pa1] **D. Alvarez-Melis** and N. Fusi. "Gradient Flows in Dataset Space". Patent 11,709,806 B2. 2022.

Theses

[T1] **D. Alvarez-Melis**. "Optimal Transport in Structured Domains: Algorithms and Applications". PhD thesis. Massachusetts Institute of Technology, 2019.

[T2] **D. Alvarez-Melis**. "The Matrix Multiplicative Weights Algorithm for Domain Adaptation". MA thesis. New York University, 2013.

[T3] **D. Alvarez-Melis**. "El Teorema de Lax Milgram, Generalizaciones y Aplicaciones". MA thesis. Instituto Tecnologico Autonomo de Mexico, 2011.

Talks

⟶ 'A Data-Centric view on Foundation Model Adaptation, Coordination and Modularity'
- 'Foundation Models in the Wild' workshop at ICML 2024, Vienna, Austria, July 2024.

⟶ 'Neural 'Surgery' with Optimal Transport: Aligning, Composing, and Repurposing Neural Networks for Modular Machine Learning'
- Optimal Transport in Cargese Workshop, Institut d'Etudes Scientifiques de Corse, Corsica, France, April 2024.

⟶ 'Data-Centric Machine Learning Methods for Heterogeneous Datasets, Domains, and Modalities'
- AstroAI Seminar, Center for Astrophysics, Harvard & Smithsonian, March 2024.

⟶ 'Machine Learning in the Space of Datasets: an Optimal Transport Perspective'
- Workshop on Applied Optimal Transport, Institute for Mathematical and Statistical Innovation, University of Chicago, May 2022.
- Topology, Geometry, and Data Analysis Seminar, Ohio State University, March 2023.
- Statistics Seminar, Harvard University, October 2023.
- Machine Learning Seminar, Center for Data Science, Boston University, November 2023.

⟶ 'Ideal made real: machine learning with limited data and interpretable outputs'
- Boston University, Faculty of Computing & Data Sciences, March 2021.
- Harvard University, Computer Science Department, February 2021.
- Northeastern University, Khorury College of Computer Science, January 2021.
- Microsoft Research New England, January 2021.
- Yale University, Department of Statistics & Data Science, January 2021.

⟶ 'Automating Dataset Comparison and Manipulation via Optimal Transport'
- Directions in Machine Learning, Microsoft, November 2020.
- Machine Learning for Data Workshop @ ICML 2021, (remote), July 2021.
- BIRS-CMO workshop on Geometry & Learning from Data, October 2021.
- AMS Spring Eastern Sectional Meeting: Special Session on Mathematics of Data Science, (remote), March 2022.

⟶ 'Geometric Dataset Distances via Optimal Transport'
- NeurIPS, (remote), December 2020.
- AutoML Workshop @ ICML, (remote), July 2020.

⟶ 'Unsupervised hierarchy matching via optimal transport'
- AISTATS, (remote), June 2020.

⟶ 'Interpretation, Representation and Correspondence in Structured Domains'
- Facebook Artificial Intelligence Research (FAIR), NYC, February 2019.
- ASAPP, NYC, February 2019.
- Google, Cambridge MA, February 2019.
- Microsoft Research, Cambridge MA, February 2019.
- IBM Research, Cambridge MA, February 2019.
- DeepMind, London, January 2019.
- Microsoft Research, NYC, January 2019.

⟶ 'Structured Optimal Transport'
- Harvard University, November 2018.
- Phillipe Rigollet's Group, MIT, November 2018.
- AISTATS, Lanzarote, April 2018.
- Optimal Transport in ML Workshop @ NIPS 2017, Long Beach, December 2017.

⟶ 'Gromov-Wasserstein Alignment of Word Embedding Spaces'
- Jim Glass's Group, MIT, November 2018
- EMNLP, Brussels, November 2018

⟶ 'Word Embeddings and Neural Networks for Natural Language Processing'
- RIIAA 2018, Mexico City, August 2018
- DeepLearn Seminar, MIT, October 2015

⟶ 'On The Robustness of Interpretability Methods'
- Workshop on Human Interpretability in Machine Learning (WHI) @ ICML 2018, Stockholm, July 2018

⟶ 'Interpretability in Natural Language Processing'
- Guest Lecture at CMU ECE-739 (remote), April 2018

⟶ 'Learning with structured data: interpretability and optimal transport'
- OpenAI, San Francisco, Januray 2018

⟶ 'Interpretability for Complex Models Natural Language Processing'
- Systems That Learn, MIT, December 2017
- CompLang Seminar, MIT, November 2017

Teaching and Mentoring

2024 **Co-Instructor**, *CS 181: Principles of Machine Learning*, Harvard University

2023 **Summer Internship Mentor**, Junhong Shen (CMU), Giorgio Giannone (Copenhagen), Neil Maillinar (UCSD), Yemin Yu (CUHK)

2022 **Summer Internship Mentor**, Jiajiao Fan (Georgia Tech), Alex Derhacobian (Stanford), Ching-Yao Chuang (MIT), Pınar Demetçi (Brown), Kianosuh Falahkheirkhah (UIUC)

2021 **IAP Micro-Internship Mentor**, Neha Hulkund (MIT)

2021 **Summer Internship Mentor**, Anna Yeaton (NYU), Wenshuo Guo (Berkeley)

2018 **Co-Supervisor, MSc Thesis**, Charlotte Bunne (MIT/ETH), Thesis award (ETH)

2017-2019 **Advisor**, Undergraduate Research Opportunities Program (5 students), MIT

Spring 2015 **Teaching Assistant**, *6.036: Introduction to Machine Learning*, MIT

Spring 2013 **Adjunct Instructor (TA)**, *MATH-UA.121: Calculus I*, NYU

Fall 2012 **Adjunct Instructor (TA)**, *MATH-UA.9: Algebra and Calculus*, NYU

Spring 2012 **Grader**, *MATH-UA.326: Analysis II*, NYU

2010 – 2011 **Teaching Assistant**, *Calculus I*, ITAM

Spring 08/09 **Teaching Assistant**, *Economics III (Intermediate Microeconomics)*, ITAM

Professional Training

June 2017 **Machine Learning Summer School**, *Max-Planck-Institut*, Tübingen, Germany

July 2014 **Regularization methods for Machine Learning**, *Univ. of Genova*, Italy

Computer Skills

Languages Python, Bash, Java, R, C++, Lua

Libraries PyTorch, Torch, Theano, Scikit

Languages

Spanish	Native	
English	Fluent	*TOEFL (iBT) 113/120, IELTS 8.5/9, FCE, CAE both with Grade A.*
Italian	Advanced	*CILS-Tre Certificate.*
French	Conversational	*Mother's language, studied also at Alliance Française Bordeaux.*
German	Basic	*Completed levels A1 - A2 at Goethe Institut Mexiko.*
Dutch, Greek	Beginner	

Professional Memberships

AMS (2012–), SIAM (2013–), ACL (2016–), AAAS (2017–), IEEE (2021–)

Other Interests

Languages, architecture, classical guitar (Albéniz, Sor), Italian cinema, soccer.