David Salinas

✉ david.salinas.pro@gmail.com
https://geoalgo.github.io/

Professional Experience

2019–Now **Senior Machine Learning Scientist at Amazon**, *Lyon*, France
- AutoML, bayesian optimization, meta-learning

2019–Now **Senior Machine Learning Scientist at NAVERLABS Europe**, *Grenoble*, France
- NLP, question answering and meta-learning

2018–2019 **Senior Machine Learning Scientist at Amazon**, *Berlin*, Germany

Developing large scale ML algorithms for applications such as forecasting and NLP
- Forecasting: proposed and implemented a deep recurrent neural network that improved state-of-the-art by 23% and got released as an AWS service in Sagemaker
- Missing value imputation: proposed and implemented a character based LSTM that achieved very high precision when imputing missing values of the Amazon catalog
- Mentoring internships and new-joiners in Machine Learning and engineering

2016–2018 **Machine Learning Scientist II at Amazon**, *Berlin*, Germany
- Developed and integrated in production a forecasting algorithm to predict labor attendance and attrition which was shown to surpass human expert accuracy.

2015–2016 **Software Development Engineer I at Amazon**, *Berlin*, Germany
- Operational and on-call tasks, development of distributed forecasting algorithms and error analysis tools for Amazon demand forecasting

2013–2015 **Post Doc at Inria**, *Sophia Antipolis*, France
- Coherent simplification of large 3D mesh of cities (≈6M vertices). Implementation in `C++`
- Active developer of `Gudhi` open-source library in `C++` (two packages

Education

2010–2013 **PhD in Computer Science at Gipsa-lab**, *Grenoble*, France
- Design and implementation of algorithms for approximating high-dimensional shapes in `C++`
- Graduate Teaching Assistant on Probability, Operating Systems, `Java` and Image Processing

2007–2010 **BSc and MSc at Ecole Normale Supérieure de Lyon**, *Lyon*, France
- Theoretical computer science degree obtained with honours

Technical skills

- Ability to design complex algorithms and data-structures for large scale problems
- Machine Learning: Deep Learning (`MXNet`, `Pytorch`), Hyperparameter Optimization, Forecasting, NLP, full ML production pipeline
- Big data: proficient with `AWS`, `SageMaker`, `Spark` and `EMR`
- Proficient with `Python`, `Scala`. Previous use of `C++`, `Java`, `OCaml`, `Matlab`

Patents

US Patent 4 US patents filed during my Amazon tenure

Publications

D. Salinas, M. Bohlke-Schneider, L. Callot, R. Medico, and J. Gasthaus. High-dimensional multivariate forecasting with low-rank gaussian copula processes. In *Neurips*. 2019.

D. Salinas, H. Shen, and V. Perrone. A copula approach for hyperparameter transfer learning. In *Advances in Neural Information Processing Systems - Metalearn workshop*, 2019.

J. Gasthaus, K. Benidis, Y. Wang, S. Sundar Rangapuram, D. Salinas, V. Flunkert, and T. Januschowski. Probabilistic forecasting with spline quantile function rnns. In *AISTATS*, 2019.

A. Alexandrov, K. Benidis, M. Bohlke-Schneider, V. Flunkert, J. Gasthaus, T. Januschowski, D. C. Maddix, S. Rangapuram, D. Salinas, J. Schulz, Lorenzo Stella, Ali Caner Türkmen, and Y. Wang. Gluonts: Probabilistic time series models in python, 2019.

F. Biessmann, T. Rukat, P. Schmidt, P. Naidu, S. Schelter, A. Taptunov, D. Lange, and D. Salinas. Datawig: Missing value imputation for tables. *Journal of Machine Learning Research*, 20(175):1–6, 2019.

D. Attali, A. Lieutier, and D. Salinas. When Convexity Helps Collapsing Complexes. In *35th International Symposium on Computational Geometry (SoCG 2019)*.

S. Schelter, F. Bießmann, T. Januschowski, D. Salinas, Stephan Seufert, and Gyuri Szarvas. On challenges in machine learning model management. *IEEE Data Eng. Bull.*, 41(4):5–15, 2018.

T. Januschowski, J. Gasthaus, Y. Wang, D. Salinas, V. Flunkert, M. Bohlke-Schneider, and Laurent Callot. Criteria for classifying forecasting methods. *International Journal of Forecasting*, 36(1):167 – 177, 2020. M4 Competition.

F. Biessmann, D. Salinas, S. Schelter, Philipp Schmidt, and D.. Lange. Deep learning for missing value imputation in tables with non-numerical data. *27th ACM International Conference on Information and Knowledge Management (CIKM 18)*, 2018.

D. Salinas, V. Flunkert, J. Gasthaus, and T. Januschowski. Deepar: Probabilistic forecasting with autoregressive recurrent networks. *International Journal of Forecasting*, 2019.

Joos-Hendrik Böse, V. Flunkert, J. Gasthaus, T. Januschowski, D. Lange, D. Salinas, S. Schelter, M. Seeger, and Y. Wang. Probabilistic demand forecasting at scale. *Proc. VLDB Endow.*, 10(12):1694–1705, August 2017.

M. W Seeger, D. Salinas, and V. Flunkert. Bayesian intermittent demand forecasting for large inventories. In *Advances in Neural Information Processing Systems 29*, pages 4646–4654. 2016.

D. Salinas, F. Lafarge, and P. Alliez. Structure-aware mesh decimation. *Computer Graphics Forum*, 34(6):211–227, 2015.

D. Attali, A. Lieutier, and D. Salinas. Vietoris-rips complexes also provide topologically correct reconstructions of sampled shapes. *Computational Geometry: Theory and Applications (CGTA)*, 46:448–465, 2012.

D. Attali, A. Lieutier, and D. Salinas. Efficient data structure for representing and simplifying simplicial complexes in high dimensions. *International Journal of Computational Geometry & Applications*, 22(04):279–303, 2012.

D. Attali, A. Lieutier, and D. Salinas. Efficient data structure for representing and simplifying simplicial complexes in high dimensions. In *Proceedings of the Twenty-seventh Annual Symposium on Computational Geometry*, SoCG '11, pages 501–509. ACM, 2011.

D. Attali, A. Lieutier, and D. Salinas. Vietoris-rips complexes also provide topologically correct reconstructions of sampled shapes. In *Proceedings of the Twenty-seventh Annual Symposium on Computational Geometry*, SoCG '11, pages 491–500. ACM, 2011.

T. Dang and D. Salinas. Image computation for polynomial dynamical systems using the bernstein expansion. In *CAV*, volume 5643 of *Lecture Notes in Computer Science*, pages 219–232. Springer, 2009.