

Dat Tran

Born April 30, 1988 in Hanoi, Vietnam

Address: Breite Straße 33
13597 Berlin, Germany

✉ E-Mail: datitran@gmail.com

☎ Phone: +(49) 151 2296-3660

🌐 Web: <http://dat-tran.com/>

🌐 LinkedIn: <https://www.linkedin.com/in/dat-tran-a1602320>

🌐 Github: <https://github.com/datitran>

🐦 Twitter: @datitran

Summary

Experienced technology and business leader with a strong focus on smart (data) products. Passion for building up high performance, cross-functional and self-managed balanced teams (from scratch). Open-source and machine learning evangelist with global reach. Currently co-heading the data team at idealo (www.ideal.de) with a special focus on building up the data science/data engineering team. My interests are diverse from traditional machine learning, deep learning, AI in general to computer vision. I have a lot of experiences from devising realistic data-driven use cases to the actual implementation into a real product; more than capable of distinguishing hype, buzzwords and wannabes from substance. Big fan of agile practices, lean product development and open-source technologies. I mainly use Python and the PyData stack for anything that I do but I also know a bit about Apache Spark, TensorFlow, R, SQL and many other stuff. I also regularly blog about my work on Medium and also speak at conferences such as PyData.

Experience

- **ideal internet GmbH** Berlin, GER
Head of Data Products at idealo *10/2017 - present*
 - Co-head of the data team (25+ people) with a special focus on building up the data science/data engineering team (built up from scratch - hired 6 people within 9 months)
 - Led the data science team directly where my team's research focus is on:
 - * Interpretability of deep learning models (mainly computer vision problems)
 - * Efficient deployment of machine learning models
 - * AutoML
 - * Data collection process (in-house labeling and external mechanical turk approaches)
 - Selected projects of my team:
 - * We used deep learning to automatically rank millions of hotel images (300+ million) according to their aesthetic and technical quality which helped to improve the listing for small and mid-sized hotels
 - * We trained CNNs to automatically generate nice looking product galleries which reduced the workflow time of our content team by 75%
 - * We built a tree-based classifier to predict non-purchases of customers which helped our CRM team to target those customers three days ahead before we receive the actual sales data by the shops
 - Provided technical and scientific guidance to senior team members and also our machine learning trainees
 - Created the machine learning powered roadmap for idealo for both product and research

- Turned the data team from a pure cost center into a profit center e.g. monetarizing our data for CRM campaigns which resulted in EUR 1 million additional revenue
- Set up a regular knowledge sharing session internally and externally across the Axel Springer companies including Bonial, Uday etc.
- Established the open-source strategy under the idealoTech umbrella and executed it e.g. created and managed the idealoTech blog (<https://medium.com/idealo-tech-blog>), idealo's Github landing page and also the @idealoTech channel on twitter
- My team open-sourced several projects with high business impact:
 - * image-quality-assessment (Convolutional Neural Networks to predict the aesthetic and technical quality of images.) - <https://github.com/idealo/image-quality-assessment>
 - * image-super-resolution (Keras implementation of the Residual Dense Network for super scaling images.) - <https://github.com/idealo/image-super-resolution>

- **Ampair** Berlin, GER
Part of the mentoring program for people in tech, <https://ampair.io/> 02/2019 - present
- **Springboard** Remote, USA
Mentor for the AI/ML career track, <https://www.springboard.com/> 01/2019 - present
- **Data Festival** Munich, GER
Member of the program committee, <https://www.datafestival.de/> 05/2018 - present
- **Thesis Advisor** Berlin, GER
Co-supervised thesis 2018
 - Hong Ngoc Truong. Bachelor Thesis, co-supervisor with Prof. Dr. Manfred-Erich Jäger-Ambrozewicz, HTW Berlin - University of Applied Sciences, Germany. Title: "Machine Learning: Tree-Based Algorithms"
- **Pivotal Inc.** Berlin, GER
Senior Data Scientist at Pivotal Labs 10/2015 - 10/2017
 - Leading and managing the Data Science practice for Pivotal Labs Berlin
 - Created machine learning models and put them into production for various clients
 - Supported pre-sales activities by creating demos/proof of concepts and blog articles
 - Scoped data science projects including proposal writing, resource estimations and budget allocation
 - Provided data science thought leadership, for example, helped to establish the smart apps idea and XP practices into data science like TDD (<http://tinyurl.com/gme62kr>), pair programming and CI/CD
 - Established an API First culture within the company: <http://tinyurl.com/z9kv5ae>
 - Gave talks at PyData Berlin, London, San Francisco 2016 and CF Summit Europe 2016
 - Data Scientist from 10/2015 - 06/2016
- **Accenture GmbH** Berlin, GER
Data Scientist at Accenture Digital 08/2014 - 09/2015
 - Built up internal capabilities and showcases in the area of network analytics, operations analytics, text analytics, customer analytics and data visualizations
 - Worked on setting up an analytics center of excellence for a large German airline
 - Led and co-ordinated workshops at customers to demonstrate the value of data science
 - Acted as point of contact for internal requests of various advanced analytics topics such as machine learning and optimization
 - Supported pre-sales activities by creating proof of concepts for clients in telecommunication, aviation and automotive

- Helped to organize internal and external events such as TEDxRheinMainSalon Datanauts and LMU DataCrunch

Sy Restaurant

Berlin, GER

- *Working in the Family Business*, <http://www.sy-restaurant.de/> 12/2006 - present
 - Supported the formation of a new branch in Berlin Spandau
 - Developed and managed the service
 - Managed online marketing activities (SEO, review portals and social networks)

Education

Humboldt University of Berlin

Berlin, GER

- *Master of Science in Business Administration, Grade 1.9* 10/2011 - 12/2013
 - Field of Specializations: Operations Research and Econometrics
 - Research Interests: Graph Theory, Queuing Theory and Dynamic Programming
 - Master Thesis: Optimal Design for an Online Dutch Auction Selling Airline Tickets
Advisor: Prof. Dr. Kurt Helmes, Department of Operations Research, Grade 1.7
Digital Copy: <http://tinyurl.com/jsq8hwp>

Berlin School of Economics and Law

Berlin, GER

- *Bachelor of Arts in Business Administration, Grade 1.5* 04/2008 - 07/2011
 - Field of Specializations: Banking & Finance and Accounting
 - Bachelor Thesis: Best of N Concept - A Comparison of Dynamic Asset Allocation Strategies
Advisor: Prof. Dr. Thomas Wolke, Department of Finance, Grade 1.7
Digital Copy: <http://tinyurl.com/184jskr>

Study Abroad at the University of Sydney

Sydney, AUS

- *Exchange Program* 03/2010 - 08/2010
 - Courses taken in Portfolio Management, International Financial Management, International Business Strategy, Calculus I and Linear Algebra I

Friedrich Anton von Heinitz Gymnasium

Rüdersdorf, GER

- *Abitur, AP English and Chemistry, Grade 2.5* 07/2007

Selected Online Publications

- **Deep Learning for Classifying Hotel Aesthetics Photos** NVIDIA Developer Blog
<https://tinyurl.com/yadn9wgk> - Jointly with Christopher Lennan 10/2018
- **What is Minimum Viable (Data) Product?** Medium - idealo Tech
<https://tinyurl.com/ybqs55bt> 06/2018
- **My two cents on what makes a good data scientist nowadays?** Medium - Towards DS
<http://tinyurl.com/y9grlbrg> 07/2017
- **8 "Simple" Guidelines For Data Projects** Medium - Built to Adapt
<http://tinyurl.com/m8nhv2o> 04/2017
- **Continuous Integration for Data Science** Pivotal Engineering Journal
<http://tinyurl.com/j6ghfht> 02/2017
- **Delivery in a Land with No Postcodes** Medium - Actionable Data
<http://tinyurl.com/gqat3bo> - Jointly with Megha Agarwal 11/2016
- **How Automakers And The IoT Are Colliding To Save Lives** Pivotal Blog
<http://tinyurl.com/jjcvaxs> 10/2016
- **Test-Driven Development for Data Science** Pivotal Engineering Journal
<http://tinyurl.com/gme62kr> - Jointly with Megha Agarwal 09/2016

- **API First for Data Science**
<http://tinyurl.com/z9kv5ae> - Jointly with Alicia Bozyk

Pivotal Engineering Journal
07/2016

Selected Talks

- **Data Natives 2018** - <https://tinyurl.com/y8mgvpsv> Berlin, GER
Demystifying the Buzz in Machine Learning! (This Time for Real) 11/2018
- **Axel Springer CTO Workshop 2018** Berlin, GER
AI Use Cases @ idealo 11/2018
- **Süddeutscher Verlag UB Fachinformation** Munich, GER
Keynote Talk: Smart Apps - The New Normal 07/2018
- **Chief Data Officer Exchange Europe 2018** Berlin, GER
Fireside Chat - Creating Profitable Data Opportunities 06/2018
- **Data Festival 2018** - <https://tinyurl.com/y7jrneej> Munich, GER
What is the big in (Big) data? And how lean can your data be... 04/2018
- **Tech Open Air / Panel Discussion** - <http://tinyurl.com/y8vngang> Berlin, GER
New Age of Analytics for Digital Products? 07/2017
- **Pivotal Big Data Roadshow** - <http://tinyurl.com/mnsgcce> Seoul, KOR
Smart data-driven apps 04/2017
- **Cloud Foundry Europe Summit 2016** - <https://youtu.be/z7AmakqpgdQ> Frankfurt, GER
Saving human lives with the IoT 09/2016
- **PyData SF 2016** - <https://youtu.be/HyjZFYDuf8> San Francisco, USA
How you really get your data science models into production - the cool way! 08/2016
- **PyData Berlin 2016** - <https://youtu.be/ZojIGRS3HLY> Berlin, GER
PySpark in Practice - Joint talk with Ronert Obst 05/2016
- **PyData London 2016** - https://youtu.be/SETpipUZ_Lc London, UK
PySpark in Practice - Joint talk with Ronert Obst 05/2016

Selected Open-Source Projects

- **Object Detector App** - <http://tinyurl.com/yyf7tmu9> ★ 970+ ♡ 500+
Real-Time Object Recognition App with Tensorflow and OpenCV
- **face2face Demo** - <http://tinyurl.com/y3u6u4cs> ★ 800+ ♡ 250+
pix2pix demo that learns from facial landmarks and translates this into a face
- **Raccoon Dataset** - <http://tinyurl.com/y29v4uhk> ★ 700+ ♡ 470+
Trained a Raccoon detector with TensorFlow's Object Detection API
- **jupyter2slides** - <http://tinyurl.com/y4qfenxv> ★ 660+ ♡ 150+
Cloud Native Presentation Slides with Jupyter Notebook + Reveal.js

Skills & Interests

- Computer
 - Very good knowledge in Python including the PyData stack: scikit-learn, pandas, matplotlib, XGBoost etc.
 - Working experiences with deep learning frameworks such as Keras and TensorFlow
 - Practical experiences in solving big data problems using Apache Spark (PySpark) and MPP databases

- Worked under a Test-Driven and CI/CD environment and constantly embraced Pair Programming
- Experiences with Microservices, PaaS and Cloud Computing (Cloud Foundry, AWS, GCP)
- Good knowledge in SQL, Git, VIM, L^AT_EX, UNIX
- Experiences with HTML, CSS and JS (mainly React)
- Industry
 - Industry experiences in Telecommunications (Postpaid, LTE and VoIP), Aviation (Customer Segmentation), Automotive (Spare parts management, Aquaplaning prediction), Banking, Logistics and E-commerce
- Languages
 - German (native), English (fluent), Vietnamese (fluent)
- Others
 - Taekwondo, Soccer (11 years Landesliga), Traveling, Literature, Cooking, History, Strategy Games
 - McKinsey Masters 2013 Regional Winner and second in the Nationals
 - e-fellows.net Scholarship

References

Dr. Michael Natusch

Global Head of Artificial Intelligence - Prudential plc
m.natusch@me.com

Holger Mai

Managing Director - Accenture Analytics
holger.mai@accenture.com



Berlin, March 10, 2019