

EXPERIENCE	<i>Senior Data Scientist</i> - DSV, Portugal	March 2023 - Now
	<ul style="list-style-type: none"><li>• Save <b>24 FTE, or €300,000 annually</b> by using Layout Large language model (LLM) to parse invoices from 43 countries into a unified data structure.</li><li>• Enhance model reliability and reduce data drift by cutting IO calls by <b>an order of magnitude of 2</b>, increasing data intake by <b>16 times</b> and enabling the <b>weekly retraining of 11 models</b>.</li><li>• Deprecate an old code base containing token classification and post-processing heuristics with a generative LLM (1.3B parameters).</li><li>• Finetune and deploy LLama3.1 in KServe inference graph to onboard new business use cases.</li></ul>	
	<i>Data Scientist</i> - Markel Insurance, USA	Sep 2022 - March 2023
	<ul style="list-style-type: none"><li>• Enable <b>customer segmentation</b> and <b>profit prediction</b> by enriching client's data with GIS data. Implement asynchronous logic, reducing the time by the <b>order of magnitude of three</b>.</li></ul>	
	<i>Data Scientist</i> - Farfetch, Portugal	Feb 2021 - Oct 2022
	<ul style="list-style-type: none"><li>• Increase the <b>GTV by \$7.4M</b> over 3 months by devising a new ranking algorithm for results of free text search queries.</li><li>• Increase the usage of distributed ML in the team by migrating the Pandas code base to Spark and providing Spark training material. <b>100% of team members</b> finish at least <b>15%</b> of the material, leading to Spark usage on a weekly basis.</li></ul>	
	<i>Computer Vision Scientist</i> - SONAH, Germany	Jan 2020 - Jan 2021
	<ul style="list-style-type: none"><li>• Increase <b>IOU from 0.3 to 0.9612</b> when detecting cars going into a parking lot by adopting and embedding a DL model to a Raspberry Pi.</li></ul>	
	<i>Junior Data Analyst</i> - InfluencerDB, Germany	Jan 2019 - Dec 2019
	<ul style="list-style-type: none"><li>• Decrease the <b>number of bugs by 5%</b> by migrating a mono-repo into micro-services architecture.</li><li>• Increase the <b>click through rate by 50%</b> by implementing and combining influencers' fakeness, growth potential, ranking and most relevant topics.</li></ul>	
PUBLICATION	<a href="#">A Machine Learning Pressure Emulator for Hydrogen Embrittlement (code)</a> - ICML SynS & ML Workshop 2023 <a href="#">Open-domain Event Extraction and Embedding for Natural Gas Market Prediction (code)</a> - CEUR Workshop, 2020	
EDUCATION	<i>Project Management</i> • Google Professional Certificate <i>AI</i> • PGM: Representation + Inference - Stanford • NLP with Attention Models + Probabilistic Models - deeplearning.ai • Mathematics for ML: Linear Algebra - Imperial College London <i>Master of Computer Science</i> • University of Bonn - Germany <i>Software Engineer (Full Scholarship)</i> • FPT University - Viet Nam	