Alex **Burlacu**

Machine Learning Engineer · Software Engineer

(+373) 69-30-53-49 | alexburlacu1996@gmail.com | alexandruburlacu.github.io | AlexandruBurlacu | alexandru-burlacu | Alex Burlacu

Skills

Machine Learning (tools)	Pytorch, Scikit-learn, MLFlow, Ray, Pandas, Numpy, Tensorflow 1.x, Keras
Infrastructure	Docker, Kubernetes, Jaeger, cron, bash with pipes and xargs, Terraform, GCP, Apache Spark
Programming	Python, Elixir, SQL, C++, played around with a few LISPs (Hy, Clojure), Haskell, Ruby, and Scala 2.x
Languages	Romanian, English, Russian. All working proficiency. Can switch rapidly between the 3.

Work Experience

Toptal

Independent ML and MLOps consultant — *May. 2022 - Present*

- I focus on end to end ML system development, MLOps, and applied Machine Learning R&D.
- I specialize on NLP projects, but can be just as proficient with CV and tabular data.
- I establish processes and build tools to accelerate the development and improvement of ML systems.
- I also directly interact with non-technical members, understanding their needs, and prioritize what needs to be done.

DevelopmentAid

Machine Learning Engineer / Team Lead (since July 2020) — *Aug. 2019 - May. 2022*

- Used machine learning (ML) and deep learning for natural language processing (NLP) on documents to make data entry more efficient.
- Developed and produced multiple ML microservices, including one to classify and tag documents through named entity recognition using PyTorch and BERT, and another to deal with an imbalanced multi-output text classification using scikit-learn.
- Defined and wrote programs for fast data annotation and synthetic data enrichment for named entity recognition (NER). Increased the dataset size from a handful of well-annotated documents to more than a hundred.
- Guided the development of new ML models and implemented practices such as ML code review, cross-validation, and replicable experiments.
- Defined some MLOps practices mainly related to model serving using Ray Serve and experiment tracking with MLflow.
- Established an observability infrastructure to reduce the number of unreported errors and accelerated bug discovery from a few days to about 10 minutes. Used Jaeger and ELK and helped in the adoption of Prometheus and Grafana.
- Defined and documented the deployment process and reduced the time to deploy trained models to less than 10 minutes. Managed a Jenkins instance and used Jenkins pipelines for that.
- Established code reviews, periodic one-on-one meetings, explicit coding best practices, and agile processes like iteration planning, planning poker, and standup meetings, reducing feature cycle time by 5x and new bugs per iteration to 0.3.
- Led a team of three junior engineers since July 2020 in developing an automated data entry solution, developing and deploying new ML models, and handling our observability and CI infrastructures.

Universite Sorbonne Paris Nord

Research Intern — *May 2021 - Oct. 2021*

- Increased sample efficiency of deep learning algorithms, mixing techniques from self-supervised, semi-supervised, and few-shot learning applicable to images and other data sources.
- Used Google Colab notebooks to run experiments, then switched to Google Cloud Platform. Provisioned with Terraform and Ansible, creating a graphics processing unit (GPU) worker and a tracking server in a single bash command within one to two minutes.
- Used MLFlow for experiment tracking and a combination of Papermill and Optuna for hyperparameter optimization.

Technical University of Moldova

University Assistant & Junior Researcher — *Sep. 2019 - Sep. 2021*

- Recreated and taught the network programming course and two lab projects focusing on concurrency primitives and networking protocols.
- Authored and lectured the real-time programming course and three lab projects covering message-based concurrency, including actor model and CSP, and message-oriented integration patterns and protocols like MQTT and XMPP.
- Overhauled and led the distributed systems and network programming courses and labs. Updated the real-time programming course and taught it as well.
- Covered diverse topics in the distributed systems course, such as data processing systems, distributed databases, microservice design patterns, and main problems of distributed systems, like the consensus, time, and exactly-once delivery.
- Mentored five final-year students for their semester project; two of them chose me as their bachelor thesis supervisor. Led labs for over 40 students per semester.

AaHa Inc.

Data Scientist (Time series forecasting) *Nov. 2018 - Feb. 2019*

- Applied statistical models, like ARIMA and ETS, and machine learning (stacked ensembles and custom neural networks using TensorFlow and Keras) algorithms for ads eCPM price forecasting.
- Provided per-project reports on experiment progress directly to executive management and was proposing the next actions.
- Developed 2 serving proof of concept, one using PMML and another using TensorFlow Serving.
- Due to qualitative changes in the data collection and vague requirements, the project was frozen.

CERN

Summer Intern *Jul. 2018 - Aug. 2018 (2 months)*

- Associate Member of the EP-SFT group. UK Science and Technology Facilities Council (STFC) grantee.
- Project Description: Benchmarking TMVA package against TensorFlow on event-by-event inference performance on multi-layered perceptrons for HEP.
- Project Purpose: Find the bottlenecks and future directions of optimization for the TMVA subpackage of the ROOT scientific package.
- Project findings: For one-by-one and small batch (< 32) inference modes, TMVA is up to 2 orders of magnitude faster than TensorFlow 1.8, built from source with AVX512 enabled, using C++ inference API.
- GitHub repo: `github.com/AlexandruBurlacu/cern-tmva-v-tensorflow-benchmark`.
- DOI: 10.13140/RG.2.2.14513.51042
- An analysis and comparative study of TensorFlow 1.x and ROOT/TMVA packages for ML inference scenarios in high-energy physics. The project aimed to understand how does TMVA compares to TensorFlow, and how do batch size, model depth, and model width affect prediction latency.
- The work was also presented as a poster at the poster session at the EEML 2019 Summer School in Bucharest.

Redox Entertainment Inc.

Machine Learning Engineer *Oct. 2017 - Sep. 2018*

- Research and Development of neural networks for Computer Vision (Medical Image Analysis of oocytes for IVF) problems. Trained over 10 bespoke neural network architectures, using unsupervised pre-training techniques like pre-training with autoencoders and even siamese networks for self-supervised learning.
- The dataset was made of a few (600 images) medical images with very low variance between positive and negative classes. Initial experiments, even with pre-trained models on ImageNet were no better than random choice. The specialized neural architecture developed by me was on par with Google's Vision AutoML, that is, state of the art on the given problem and data, achieving 68.7% accuracy.
- Debugged a data preprocessing issue which was leaking the test set and wrongfully giving very high accuracy during evaluation. Prevented releasing the broken model, thus saving the company's reputation.
- Directly reported to the company director and was always involved in discussions about ways to increase the performance of our neural networks with invited experts, mostly PhDs.
- Mentored and trained a PhD intern for 3 months.

FAF NGO

Co-Founder and Corporate & Financial Responsible *Sep 2017 - Oct. 2018*

- Starting May 2017 I was leading the team responsible for founding the NGO, prepared all documentation required at the ministry of justice for registration, drafted the internal code of conduct and the long-term strategic plan.
- Starting September 2017 was assigned as the Corporate and Financial Responsible, leading the efforts for finding long- and short-term partners for the non-profit.
- Defined 2 budget templates, for an open lectures series, and hackathons.
- Also proposed and supported a new annual hackathon and a student internship program.

BookVoyager (Startup)

Co-Founder and CTO *May 2017 - Sep. 2017*

- BookVoyager was a recommendation/search system for fiction books that extracts features from raw text and based on these gives a recommendation.
- As the CTO I was responsible for defining the technological stack and the architecture of the solution.
- Chose Mongo as the operational database, designed an API service written in Flask and a custom search engine, communicating via XML-RPC. Also set up logging to allow for easier troubleshooting.
- Built the feature extraction and recommendation sub-systems based on token-level and whole-text analysis with SpaCy.
- Sped up the computation of recommendation results 85x by using a pre-allocated array and used profiling to identify the bottleneck.
- Was actively involved in customer interviews, defining business and development processes and was pitching the project at various venues.
- The project was named the best startup of the Founder Institute Chisinau 2017 semester.

Education

Google Cloud

Google Cloud Certified Professional Cloud Architect — *Dec. 2022 - Dec. 2024*

- Certificate ID number: **1iXGX7**

Google Cloud

Google Cloud Certified Professional Machine Learning Engineer — *Dec. 2022 - Dec. 2024*

- Certificate ID number: **K1fhzJ**

Cloud Native Computing Foundation

Certified Kubernetes Application Developer (CKAD) — *Jun. 2022 - Jul. 2025*

- Certificate ID number: **LF-7q5ie3w6bc**

Joint degree with UTM(Technical University of Moldova) and USV(Ștefan cel Mare University Suceava)

Chișinău, Rep. of Moldova; Suceava, Romania

M.E. in Computer Science and Engineering — *Sep. 2019 - Sep. 2021*

- Researched and developed a system combining information retrieval, self-supervised, and few-shot learning techniques to achieve well-performing deep learning models with few samples.
- Defended the Master thesis at the Technical University of Moldova among the top 3 students in our batch, in December 2020, and at Ștefan cel Mare University Suceava with the maximum grade in September 2021.
- Successfully completed courses on Cloud Computing and Infrastructure, Advanced Network Security, and a few other ones at UTM. Also, among other lecture series, attended HCI and Augmented Reality, and a Distributed Computing ones at USV.

EEML Summer School

Bucharest, Romania

Atendee — *Jul. 2019 - Jul. 2019*

- One week summer school organized by the Technical University of Bucharest and DeepMind
- Got familiar with self-supervised learning techniques for vision, basic RL like SARSA and Q-Learning.
- Presented my work done at CERN at the poster session.

Aalborg Universitet

Aalborg, Denmark

B.S in Robotics — *Jan. 2019 - Jun. 2019*

- Erasmus+ student exchange program.
- Got familiar with the Project Based Learning paradigm, fundamentals of Computer Vision and Robotics.

UTM(Technical University of Moldova)

Chișinău, Rep. of Moldova

B.E. in Computer Science and Engineering — *Sep. 2015 - Jun. 2019*

- Studied at the English Taught Honors Program in Computer Science.

Other Activities

Toptal

Remote

Technical Screener — *Aug. 2022 - Present*

- Conducted over 60 technical interviews for Data Science and Artificial Intelligence specializations.
- Evaluating take-home assignments by candidates
- Creating new interview tasks

"Designing a Serverless Platform" Course

Remote

Independent Instructor — *Oct. 2021 - Jan. 2022*

- Created a course about the inner workings and the system design of serverless platforms.
- Other than the main topic, serverless computing, the course also covers such topics as developer experience (DX), containerization technologies, distributed systems and software scalability and optimizations.
- The course is a combination of live coding and whiteboard explanations.
- The first edition of the course is free and invite-only; 24 people were invited, with an attendance rate per lecture of over 50%
- Code available at `https://github.com/AlexandruBurlacu/ServerlessCourseAutumn2021`.

Ministry of Education and Research of Republic of Moldova/TwentyTu

Chișinău, Rep. of Moldova

Co-Author/Group Coordinator — *Jan. 2020 - March. 2021*

- Lead the working group for the development of the Artificial Intelligence optional course curriculum for high-school students. `https://mecc.gov.md/sites/default/files/curriculum_ia_aprobat_cnc.pdf`
- Author of the Artificial Intelligence-related chapters (1, 3, 4) of the optional course, spanning such topics as AI applications, simple and advanced neural networks for computer vision and natural language processing, classic ML algorithms, data cleaning and analysis, linear algebra, and non-technical topics such as: AI bias and fairness, missuses of AI, AI from a legal perspective.
- The course contains both textbook-like materials and a set of interactive exercises in the form of Jupyter Notebooks.