Lucian Istrati

Curriculum Vitae

Street "Nicolae Titulescu" nr. 39, Sector 1, Bucharest
011132
Romania
+40 762 048 376
lucianistrati@yahoo.com

Education

2018–2021 **Faculty**, *University of Bucharest, Faculty of Mathematics and Computer Science*, Bucharest, Ranked 16th out of 1019 candidates in the admission contest, student in the "Performance Group", Weighted average mark - 8.92, BSc. thesis mark - 10.00.
Undegraduate field: Computer Science

2021–2023 **Master**, *University of Bucharest, Faculty of Mathematics and Computer Science*, Bucharest, M. Sc. in Data Science.
Ranked 1st out of 44 candidates in the admission contest, Admission mark - 9.84

Work experience

- **AI Scientist I** at Ordaos Bio (Oct 2021 - Present)
- **Machine Learning Engineer** at EmailTreeAI - worked on a Named Entity Recognition pipeline and an Information Extraction component; used NLTK, SpaCy, Flair, Torch, Tornado (Aug 2021 - Sep 2021)
- **Machine Learning Engineer** at Metaminds - worked on multiple Computer Vision projects, used OpenCV, PyTorch, Imgaug, Numpy, Matplotlib (Nov 2020 - Jul 2021)
- **Junior Machine Learning Intern** at Metaminds - worked on a Computer Vision project; used OpenCV, Keras, Numpy, Matplotlib (Oct 2020)
- **Junior Machine Learning Intern** at Romanian Intelligence Service(RIS) - worked on a Natural Language Processing project; used Keras, Pandas, Numpy, Matplotlib, Sklearn, NLTK, Blob, SpaCy, Fasttext, Flask (Jun 2020 - Sep 2020)
- Teaching "Programming of Algorithms" and "Data Structures" as a tutor for students in their first year of Undergraduate Studies of Computer Science at Faculty of Mathematics and Computer Science, University of Bucharest. (Oct 2019 - May 2020)

Research experience

- First running author of a published paper entitled **"Automatic monitoring and analysis of brands using data extracted from Twitter in Romanian"**. The paper has been accepted at IntelliSys 2021 Conference and will be published in the Springer series "Advances in Intelligent Systems and Computing"

Programming languages and Technologies

Experimented Python, Numpy, Pandas, Sklearn, Keras, PyTorch, C++

Advanced OpenCV, NLTK, SpaCy, Imgaug

Average	SQL	Begginer	R, Flask, Tornado

Courses

- Enrolled in the Udemy Course "Complete Data Science Bootcamp" provided by 365 Careers
- Completed Udemy Course "Natural Language Processing: NLP With Transformers in Python"
- "Machine Learning Fundamentals CS302" - Machine Learning Course in Python - TelAcad. (Jul 2019 - Sep 2019)

Contests

- "Word Complexity Estimation Predict the probability of a word being complex" Kaggle Competition - ranked 13/89 (Nov 2021 - Dec 2021)
- SEMEVAL 2022 - "Multimedia Automatic Misogyny Identification" Competition - ranked 32/83 on Task B and 52/83 on Task A (Jan 2022)
- Textract NLP Hackathon - Classifying news articles in either automatically generated or written by human authors. (07.12.2019 - 08.12.2019)

Projects

- Involved in organizing a machine learning hackathon named "Nitro Language Processing Hackathon". Choosed Named Entity Recognition for the hackathon task after performing several experiments. Scrapped several datasources in order to create a novel corpus which was then adnotated with named entities by a team of 12 students. Created the Kaggle Competition itself in which 24 teams from different universities and highschools competed (Feb 2022 - March 2022)
- Worked on a data analysis & visualization team project based on Beverages & Breweries dataset - used several clustering and dimensionality reduction techniques
- Worked on a "WhatsApp Group Chat Data Analysis" team project involving Text Classification with Curriculum Learning methods (Dec 2021 - Jan 2022)
- Developed a Stock Price prediction series of backtested models that use Deep Learning and NLP techniques / algorithms in order to trade stocks based previous stock prices, financial reports and news analysis (Jan 2021 - May 2021)
- Developed an adapted algorithm for topic modelling that uses Sentence Embeddings in order to filter out noisy clusters (Apr 2021 - May 2021)
- Worked on a team project on "Hate Speech Author Profilling" with an ablation study comprised of several DL arhitectures (Feb 2021 - Mar 2021)
- Built a CoReference Resolution ML model where NLP features are derived from a text and fed as input to a model able to predict whether a pair of tokens is a coreference or not (May 2021)

- Worked in a Computer Vision team project for the "University of Bucharest Software Development Center" using Handwritten Recognition and Optical Character Recognition techniques with the aim to extract data from the documents highschool students send in order to participate in the Admission Contest of the Faculty of Mathematics and Computer Science, University of Bucharest. (Mar 2020 - Jun 2020)
- Human organ segmentation from images in DICOM format using Computer Vision and different Clustering Algorithms (Mar 2020)
- Binary classification of hate/non-hate tweets and topic prediction for StackOverflow Questions using Python and Machine Learning (Sep 2019)