Santiago **Dubov**

ENGINEERING MASTERS STUDENT AT THE UNIVERSITY OF CAMBRIDGE WITH STRONG ANALYTICAL AND PROBLEM SOLVING SKILLS. HIGHLY INTERESTED IN MACHINE LEARNING AND APPLIED SCIENCE

Corpus Christi College, Trumpington Street, Cambridge, CB2 1RH

☐ +44 7429955869 | ✉ sd793@cam.ac.uk | in santiago-dubov

Education

University of Cambridge
Cambridge, UK

M.ENG ENGINEERING (4TH YEAR), SPECIALISATION: COMPUTER ENGINEERING AND MACHINE LEARNING *oct. 2017 - june 2021*

- 1st Year,2nd Year,3rd Year: Class I(6th percentile), Class I (12th percentile), Equivalent Class I
- Third year completed on exchange with CentraleSupélec.
- Fourth year Masters project representing 50% of final grade: Automated grammar correction of non-native english speech using Deep Neural Networks and NLP techniques.
- Relevant Courses: Probabilistic Machine learning, Deep learning, Computational Statistics, Software Engineering, Optimisation.

Université Paris-Saclay: CentraleSupélec
Paris, France

INFORMATION ENGINEERING SPECIALISATION, EXCHANGE YEAR (ERASMUS) *August 2019 - July 2020*

- Courses: Artificial Intelligence, Machine Learning, High Performance Computing, Cloud Computing, Optimisation, Software Engineering (Java), 8 month technical project in a group of 5.

Park View School
County Durham

A LEVELS *2010 - 2017*

- 3 A*: Maths, Physics, Chemistry, 2 A at AS level: Further Maths, French

Skills

Programming	Python, JAVA, C++, Julia, LaTeX, UNIX, MYSQL, Matlab
Software	Pandas, Seaborn, SKlearn, PyTorch, Tensorflow, Docker
Languages	English (native), French (C1), Italian (C1), Portuguese (B2), Russian (B1)

Work Experience and Academic Projects

Spotify
London

MACHINE LEARNING ENGINEER INTERN *Jun. 2021 - Sep 2021*

- Working with the Search team to improve discoverability and reduce bias in search result rankings.

Improving Automated Spoken Grammar Correction
University of Cambridge

FINAL YEAR MASTERS PROJECT *Aug. 2020 - June 2021*

- Used sequence to sequence models (Transformers) to convert between ASR-transcribed grammatically correct and incorrect text.
- Pre-processing of corpora and data augmentation for low resource tasks.
- Used masked language models (BERT) to introduce speech disfluencies in augmented data.
- Constructed language models such as n-grams for filtering of over-generated pseudo data.

Foreign Object Detection on Runways
CentraleSupélec

PROJECT SOFTWARE TEAM LEAD (TEAM OF 6) *Sep. 2019 - July 2020*

- Developed a new method for the detection of foreign objects on runways using a Laser and computer vision (OpenCV).
- Won 1000€ and second prize in the Aerosaclay Airport of the Future competition.
- Trained neural networks using transfer learning in TensorFlow to compliment laser detection and provide additional security to the detection system.

Faraday Predictive (Start-Up)
Cambridge, UK

DATA SCIENTIST/SOFTWARE DEVELOPER *Jun. 2018 - Aug. 2018*

- Performed experimental analysis of mechanical faults on an Induction motor test rig to improve predictive maintenance systems.
- Analysed test results and customer data from SQL using software tools created in python. Data visualisation with Seaborn.
- Compared the companies hardware and software to those of competitors to produce a product analysis.

Extracurricular Activities

Rowing
CORPUS CHRISTI COLLEGE BOATCLUB AND AVIRON CENTRALESUPÉLEC PARIS

Cambridge/Paris

Nov. 2017 - Present

- High volume of training (around 8 times per week).
- Competed in national and international competitions.

International Board of CentraleSupélec
TRAVEL TEAM LEADER

CentraleSupélec

Nov. 2019- June 2020

- Involved in the integration of approximately 100 international students to the campus.
- Organisation of trips and events to promote cultural integration.

Personal Projects
COMPLETED IN FREE TIME

2020

- Followed various NLP and AI courses: UC Berkely, Standford, Coursera.
- Kaggle and hackathon enthusiast.
- Created a website detailing the most effective way of learning foreign languages with links to the best resources.