Karolis Jankauskas

A Biochemical Engineer with research experience turned Data Scientist.

Experienced in building and scaling deep learning solutions, backend development, statistical analysis, and operational research.

Languages Python • Cython • C++ • CUDA • SQL • R • Julia
Libraries Keras • TensorFlow • PySpark • Xgboost • rpy2 • Plotly • Dash • OpenCV
Backend Flask • Django • Celery • Redis • RabbitMQ • Docker • Kubernetes
Databases PostgreSQL • MySQL • InfluxDB
Other CircleCI • AWS

London, UK

+44 7858 271487
info@kjankauskas.com

www.kjankauskas.com
www.github.com/karolisjan
www.linkedin.com/in/karolis-jan

EXPERIENCE

Data Engineer
WorldQuant
London, UK, 2018 September – Present

WorldQuant is a quantitative asset management firm with more than 700 employees across 26 offices in 15 countries.

- Developed an encoder-decoder LSTM-based multiple time series anomalies detection app with a dashboard and REST API endpoints.

Machine Learning Engineer
Aiden.ai
London, UK, 2018 February – 2018 September

Aiden is a Natural Language Processing (NLP) powered virtual assistant who helps marketers make better decisions.

- Developed a dockerised REST API app to detect anomalies in advertising time series data in real-time.
- Built a multi-task deep neural network-based model with entity embeddings of categorical variables to answer questions such as "*What will the total number of impressions, link clicks, and app installs be next week if I spend $X on ads ABC targeting Y audience in Z location?*".
- Created a neural network and genetic algorithm-based system for recommending an optimal way to distribute weekly advertising budget to maximise ad sets performance, e.g. impressions, link clicks, and app installs.

Data Scientist
Picasso Labs
London, UK, 2017 Feb – 2018 February

Picasso Labs is recognised as one of Unilever Foundry's most ambitious and innovative start-ups of the past 5 years.

- Using a pre-trained Inception v3 CNN, HSV color histograms, and entity embeddings of date, time, and image tags developed a system for recommending images to improve marketing performance on social media.
- Re-used the system above to create a content-based image retrieval engine.
- Built a custom facial expressions recognition model and applied it on over 5000 web-scraped images from US online media to investigate "visual bias". See www.newsweek.com/liberal-media-not-biased-trump-thinks-703291
- Applied regression and non-parametric statistical tests to determine best performing image categories and segments.

Teaching Assistant
UCL Biochemical Engineering
London, UK, 2016 February – 2018 February

Supervised research and taught MSc and MEng students discrete-event simulation, mathematical programming, evolutionary programming, and multi-objective optimisation.

Consultant (internship)
Sphere Fluidics Ltd
Cambridge, UK, 2014 June – 2014 August

Created fluid-flow models of microfluidic chips for a novel single-cell screening and analysis system.

Research Associate
UCL Advanced Centre for Biochemical Engineering
London, UK, 2013 June – 2013 August

Performed multi-variate data analysis on mass spectrometry data to improve the expression of virus-like particles from *Pichia pastoris* cells for a universal influenza vaccine project.

EDUCATION

Udacity
2017 – 2018
Nanodegree, Artificial Intelligence, Certificate of Completion

UCL
London, UK, 2014 – 2018
PhD, Biochemical Engineering (Operational Research)

Thesis title: "Biopharmaceutical Capacity Planning using a Flexible Genetic Algorithm Approach"

Using Python, C++, CUDA, and Docker, developed a cross-platform genetic algorithm-based tool for continuous-time multi-objective planning and scheduling of biopharmaceutical facilities under deterministic and stochastic demand.

Accomplishments

- Presented a keynote lecture at the 27th European Symposium on Computer Aided Process Engineering (ESCAPE 27), Barcelona, Spain, 2017**.**
- Awarded a Year 1 Research Project Prize for Best PhD Project and Poster.

UCL
London, UK, 2010 – 2014
Master of Engineering (MEng), Biochemical Engineering, First-Class Honors

Accomplishments

- Received Jacobs Engineering Design Project Prize.
- Received Head of Department Commendation Award.

Activities

- Fitness instructor at UCLU Muay Thai Club.

PUBLICATIONS & TALKS

Jankauskas K, Papageorgiou, L.G. and Farid, S.S., Fast Genetic Algorithm Approaches to Solving Discrete-Time Mixed Integer Linear Programming Problems of Capacity Planning and Scheduling of Biopharmaceutical Manufacture, Computers and Chemical Engineering (2018), DOI: https://doi.org/10.1016/j.compchemeng.2018.09.019

Jankauskas, K., Papageorgiou, L.G. and Farid, S.S., 2017. Continuous-Time Heuristic Model for Medium-Term Capacity Planning of a Multi-Suite, Multi-Product Biopharmaceutical Facility. In *Computer Aided Chemical Engineering* (Vol. 40, pp. 1303-1308). Elsevier, DOI: https://doi.org/10.1016/B978-0-444-63965-3.50219-1

Jankauskas, K., McCartney, GR., Osborne, MD., Papageorgiou, LG., Farid, SS. 2017. Multi-Objective Capacity Planning for Multi-Product Biopharmaceutical Facilities Under Uncertainty, 253rd ACS National Meeting, San Francisco, USA, April 2-6.

Jankauskas, K., Papageorgiou, LG., Farid, SS. 2016. Production Scheduling of a Multi-Product Biopharmaceutical Facility Using a Genetic Algorithm, 28th European Conference on Operational Research (EURO), Poznan, Poland, July 4-8.

PROJECTS

- www.github.com/karolisjan/BiopharmaScheduling/tree/development a scheduling library for biopharmaceutical facilities
- www.github.com/karolisjan/DeepLearning a collection of deep learning projects
- www.github.com/karolisjan/Genetic-Programming artificial ant and snake game agents created using genetic programming