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
Databases PostgreSQL • MySQL • InfluxDB
Other Continuous Integration • AWS • Docker • Kubernetes

London, UK

+44 7858 271487
info@kjankauskas.com

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

EXPERIENCE

Data Engineer (internship)
WorldQuant
London, UK, 2018 September – Present

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

- Using Dash, Celery, Redis, and TensorFlow, developed an encoder-decoder LSTM-based anomalies detection app with a dashboard and REST API endpoints. Deployed it in a Kubernetes cluster.

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.

- Using Flask, Python/R, and Docker, developed and deployed a REST API app to detect anomalies in advertising time series data.
- Using Keras, built a 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 ad sets targeting Y audience in Z location?*".
- Created a neural network and genetic algorithm-based approach for recommending an optimal way to distribute the weekly advertising budget to maximise the return on ad spend (ROAS).
- Using Flask, Celery, and Redis, created scheduled parallel data download and processing pipelines.

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 Keras and OpenCV, developed a convolutional neural network (CNN)-based model with additional input branches for HSV color histograms and entity embeddings of date, time, and image tags for recommending images to improve CTR or engagement on social media.
- Re-used the model above to create a content-based image retrieval (CBIR) engine.
- Using Keras and OpenCV, 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 non-parametric statistical tests to determine best performing image categories.

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++, and CUDA, developed a cross-platform genetic algorithm and Monte Carlo simulation-based tool for continuous-time multi-objective planning and scheduling of biopharmaceutical facilities with uncertain product 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 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