Detection of new insecticidal genes using ML

Dr. Asaf Levy - The Hebrew University of Jerusalem

Dr. Asaf Levy - The Hebrew University of Jerusalem

We are interested in deciphering the function of genes involved in interactions between bacteria and other bacteria, plants, and insects.

www.asaflevylab.com

alevy@mail.huji.ac.il

What is the project's research question?

Develop a machine learning classifier with our group members.

What data will be worked on?

Genes from bacterial genomes that are publicly available.

What tasks will this project involve?

Large scale data genomic data analysis, feature extraction, reading some scientific papers.

What makes this project interesting to work on?

It is important for having a sustainable agriculture. Plants engineered with insecticidal genes are protected from pests (read about Bt corn for example).

What is the expected outcome?

A list of genes that we will validate in experiments in the lab.

What infrastructure, programs and tools will be used? Can they be used remotely?

Whatever you want. Usually, we use Python programming in the group. Can be used remotely.

Is the project open source?

No

What skills are necessary for this project?

Programming, data analysis, some experience in machine learning.

Interested candidates should be at Master, PhD, or Postdoc level.

Using AI tools to determine reading difficulties.

Dr. Tzipi Horowitz Kraus - Technion, Dept. of Education in Science and Technology

Dr. Tzipi Horowitz Kraus - Technion, Dept. of Education in Science and Technology

We are using neuroimaging data to determine neuronal patterns associated with language and reading impairments in children. We use different data modalities, ranging from fMRI, DTI, anatomy, EEG, fNIRS to reveal specific patterns associated with developmental changes related to communication and AI tools for prediction or classifications related to different profiles and response to treatments.

https://www.technion.ac.il

tzipi.kraus@technion.ac.il

What is the project's research question?

Can we detect specific brain patterns associated with positive response to treatment?

What data will be worked on?

fMRI

What tasks will this project involve?

Data pre and post processing, computational models.

What makes this project interesting to work on?

The transformative nature of this project, the convergence of state-of-the-art neuroimaging data analysis tools combined with AI will make this project an intriguing one for students who are interested in biomedical/computational/data science project.

What is the expected outcome?

A scientific paper.

What infrastructure, programs and tools will be used? Can they be used remotely?

Python and Matlab. Cannot be used remotely.

Is the project open source?

No

What skills are necessary for this project?

Scripting is an advantage.

Interested candidates should be at PhD level.

Neural encoding of natural stimuli - an fmri student.

Dr. Yossi Yovel - Tel Aviv, Neuroscience.

Dr. Yossi Yovel - Tel Aviv, Neuroscience.

Animal behaviour and decision making.

www.yossiyovel.com

yossiyovel@gmail.com

What is the project's research question?

How does the brain encode natural stimuli?

What data will be worked on?

Fmri images

What tasks will this project involve?

Connecting the MRI signal to the movie observed by the scanned humans.

What makes this project interesting to work on?

It touches on one of the most fundamental questions in neuroscience.

What is the expected outcome?

A model connecting the natural stimulus and the brain.

What infrastructure, programs and tools will be used? Can they be used remotely?

Matlab or Python. Can be used remotely.

Is the project open source?

Can be

What skills are necessary for this project?

Programming, ML.

Interested candidates should be at Master level.

Markov chain-based analysis of air-pollution and climatologic systems.

Dr Barak Fishbain - Technion - Israel Institute of Technology

Dr Barak Fishbain - Technion - Israel Institute of Technology

The Technion Enviromatics Lab (TechEL) focuses on Enviromatics, a new research field that aims at devising machine learning methods and mathematical models for better understanding built and natural complex environments. The goal is to harness new machine learning, mathematical models with engineering principles, computing, and networked sensing data analytics for enhancing the efficiency, resiliency, and sustainability of infrastructure and natural systems. This includes topics related to hydro-informatics, atmospheric-informatics, traffic data, structural health, smart infrastructure systems and connected transportation.

https://fishbain.net.technion.ac.il/

fishbain@technion.ac.il

What is the project's research question?

Data science approach for inferring which affects air-pollution more meteorology or anthropogenic activities?

What data will be worked on?

Climatologic and air-pollution datasets.

What tasks will this project involve?

Developing Markov-chain based model to infer the temporal behavior of meteorological and air-pollution systems and infer the effect of the former on the latter.

What makes this project interesting to work on?

Developing mathematical models for physical phenomena and inferring their physical properties through math.

What is the expected outcome?

Developing mathematical models for physical phenomena and inferring their physical properties through math.

What infrastructure, programs and tools will be used? Can they be used remotely?

The TechEL holds a virtual computation and storage platform (VMWare based) as well as Azure (Windows based) processing resources. These will be used for data storage processing with any software the student prefers (Python, Matlab, R, etc.). Given access over VPN to

the Technion's intranet - all resources are available. Access can be granted through formal procedure.

Is the project open source?

Yes

What skills are necessary for this project?

Mathematical thinking, coding skills.

Interested candidates should be at Master, PhD or, Postdoc level.

Robust Optimization of Water Supply Systems.

Dr. Barak Fishbain - Technion - Israel Institute of Technology

Dr. Barak Fishbain - Technion - Israel Institute of Technology

The Technion Enviromatics Lab (TechEL) focuses on Enviromatics, a new research field that aims at devising machine learning methods and mathematical models for better understanding built and natural complex environments. The goal is to harness new machine learning, mathematical models with engineering principles, computing, and networked sensing data analytics for enhancing the efficiency, resiliency, and sustainability of infrastructure and natural systems. This includes topics related to hydro-informatics, atmospheric-informatics, traffic data, structural health, smart infrastructure systems and connected transportation.

https://fishbain.net.technion.ac.il

fishbain@technion.ac.il

What is the project's research question?

Robust optimization has been out there for quite some time. So far, when it comes to water network, robust optimization has been applied solely to the supply side and in the design of the network phase. Here we opt to utilize robust optimization to the operational aspects of the network.

What data will be worked on?

Simulated data, generated by EPA-NET.

What tasks will this project involve?

Developing the optimization algorithms and coding them, working with water supply system simulation software (EPA-NET).

What makes this project interesting to work on?

Robust optimization is a strong tool that can be used in many applications and thus learning it, provides significant skills toolset.

What is the expected outcome?

A research paper describing for the first time operational ROBUST optimization of water systems.

What infrastructure, programs and tools will be used? Can they be used remotely?

The TechEL holds a virtual storage and computation system, which will provide the infrastructure of the project. Given access to the Technion's intranet all resources are available. Access can be granted through formal procedure.

Is the project open source?

No

What skills are necessary for this project?

Knowledge (and preferably experience) in optimization, coding (Python or R).

Interested candidates should be at PhD or Postdoc level.

Deep learning in genomics.

Dr. Yaron Orenstein - Bar-Ilan University, Computer Science and Life Sciences.

Dr. Yaron Orenstein - Bar-Ilan University, Computer Science and Life Sciences.

Our lab develops algorithms to infer predictive models of molecular interactions based on high-throughput biological data.

https://wwwee.ee.bgu.ac.il/~cb/index.html

yaron.orenstein@biu.ac.il

What is the project's research question?

We will take a biological phenomena on the molecular level, which has been measured in high-throughput, and develop, train, and test a deep neural network on the measurements. We will desire high prediction performance and interrogate the network for the molecular principles it learned.

What data will be worked on?

There are many publicly available high-throughput genomic datasets, which are ideal for machine learning and specifically deep learning.

What tasks will this project involve?

Developing the network analytically, programming in python to implement the network, running the code on available data to train and test the network, and applying computational techniques to interrogate the network.

What makes this project interesting to work on?

High-throughput genomic data is ideal for deep neural network - there are many data points, and each data point is relatively small compared to vision tasks. There's also the scientific discovery in the step of interpreting the trained network.

What is the expected outcome?

A predictor with high prediction performance, and interpretation of the network. These may lead to a publication.

What infrastructure, programs and tools will be used? Can they be used remotely?

Unix OS, Python, Keras and Tensorflow or Pytorch. Can be used remotely.

Is the project open source?

Yes

What skills are necessary for this project?

Mathematics background (linear algebra, calculus, statistics, probability), programming, algorithms, data structure, computer architecture.

Interested candidates should be at Master, PhD, or Postdoc level.

Three-dimensional scene reconstruction using echolocation echoes.

Dr. Igal Bilik - Ben Gurion University, School of Electrical and Computer Engineering.

Dr. Igal Bilik - Ben Gurion University, School of Electrical and Computer Engineering.

Statistical and machine learning-based signal processing for sensors and sensor arrays enabling autonomous platforms operation. Smart Sensing Lab research focuses on signal processing for radar, visual and acoustic sensors data.

https://danielbilik2003.wixsite.com/igalbilik

bilik@bgu.ac.il

What is the project's research question?

Can artificial neural network-based approach perform 3D data reconstraction using echolocation echoes?

What data will be worked on?

Database of 623K echoes simulated and recorded using bio-mimetic sonar system.

What tasks will this project involve?

Derive, implement, and test the Deep neural network-based approach for the 3D scene reconstraction using the dataset.

What makes this project interesting to work on?

There does not exist efficient approach for 3D scene reconstraction.

What is the expected outcome?

Journal publication summarizing hte derived DNN-based approach for 3D scene reconstruction.

What infrastructure, programs and tools will be used? Can they be used remotely?

Computer with GPUs, Python with NN libraries. Can be used remotely.

Is the project open source?

No

What skills are necessary for this project?

Deep Learning, Statistical Signal Processing, Acoustic signal processing.

Interested candidates should be at Master, PhD or, Postdoc level

Detection, Localization, and classification of human operating in adjacency with robotic arm.

Dr. Igal Bilik - Ben Gurion University. School of Electrical and Computer

Dr. Igal Bilik - Ben Gurion University. School of Electrical and Computer

Statistical and machine learning-based signal processing for sensors and sensor arrays enabling autonomous platforms operation. Smart Sensing Lab research focuses on signal processing for radar, visual and acoustic sensors data.

https://danielbilik2003.wixsite.com/igalbilik

bilik@bgu.ac.il

What is the project's research question?

Detection, Localization, and classification of human operating in adjacency with robotic arm.

What data will be worked on?

Vision and radar recordings of human and robotic arm operating in laboratory environment.

What tasks will this project involve?

Development of Deep learning-based approach for detection, localization, and classification of human and robotic arm.

What makes this project interesting to work on?

Ability to localize and classify between human and robotic arm is critical to introduction of the consumer cobots into human living environment. Reliable algorithms are still missing, and this project can provide a basis for the new era of robot-human interaction. The project can have significant academic and industrial implications.

What is the expected outcome?

Journal article summaries derived algorithms and presenting results.

What infrastructure, programs and tools will be used? Can they be used remotely?

Computer with GPUs , Python with DLL libraries. Can be used remotely.

Is the project open source?

No

What skills are necessary for this project?

Statistical Signal Processing, Computer vision, Deep learning.

Interested candidates should be at Master, PhD or, Postdoc level

Computational disentanglement of single-cell data.

Dr. Mor Nitzan - School of Computer Science and Engineering, The Hebrew University of Jerusalem.

Dr. Mor Nitzan - School of Computer Science and Engineering, The Hebrew University of Jerusalem.

Our research is at the interface of Computer Science, Physics, and Biology, focusing on the representation, inference, and design of multicellular systems. We develop computational frameworks, based on ideas rooted in dynamical systems theory and machine learning to better understand how cells encode multiple layers of spatial and temporal information, and how to efficiently decode that information from single-cell data. We aim to uncover organization principles underlying information processing, division of labour, collective cellular function, and self-organization of multicellular structures.

https://nitzanlab.com/

mor.nitzan@mail.huji.ac.il

What is the project's research question?

Can we tease apart distinct biological processes underlying the state of cells based on single-cell data.

What data will be worked on?

Publicly available single-cell datasets.

What tasks will this project involve?

Depending on the student's/postdoc's background and interests, the project can involve a subset of the following: method development, modeling and simulations, single-cell data analysis, and biological interpretation.

What makes this project interesting to work on?

This is an exciting, emerging direction in single-cell analysis, rooted in the realization that cells simultaneously encode in their transcriptome multiple layers of information about their collective physical configuration in the tissue-of-origin, temporal processes such as the cell cycle and differentiation, and response to external stimuli. The possibility to disentangle and manipulate these different layers of information is expected to deepen our understanding of collective behaviour, information encoding and transfer, and division of labour in diverse biological systems.

What is the expected outcome?

Contribution to research paper, Contribution to software development.

What infrastructure, programs and tools will be used? Can they be used remotely?
Coding in Python, access to university servers. Can be used remotely.

Is the project open source?
Yes

What skills are necessary for this project?
A computational/mathematical background (CS/Physics/Applied Math/similar).

Interested candidates should be at PhD or Postdoc level.

Boosting Quadrotor Navigation Using Data-Science in GNSS-Denied Environments.

Dr. Itzik Klein - University of Haifa, the Autonomous Navigation and Sensor Fusion Lab.

Dr. Itzik Klein - University of Haifa, the Autonomous Navigation and Sensor Fusion Lab.

The purpose of navigation is to determine the position, velocity and attitude of platforms, humans, and animals. Obtaining accurate navigation commonly requires fusion between several sensors. The Autonomous Navigation and Sensor Fusion Lab (ANSFL) vision is augmenting and developing artificial intelligence (AI) algorithms in innovative breakthrough research to create meaningful knowledge to the society through collaboration with fellow researchers and engineers. Our pioneering research addresses the intersection of AI with the navigation and inertial sensors to create value and opportunities for ocean and environment protection, identifying illnesses and well-being in humans and animals, and developing tools for autonomous vehicles teamwork.

We have ongoing projects on AI navigation and sensor fusion, inertial sensing, pedestrian navigation, animal navigation and localization, drone sensor fusion, autonomous underwater vehicle navigation, mobile robot navigation, and much more

What is the project's research question?

How data-science can help improve quadrotor navigation in GNSS-denied environments.

What data will be worked on?

Inertial sensors (accelerometers, gyroscopes, magnetometers, barometer) - time series.

What tasks will this project involve?

Data exploration, network architecture derivation, performance evaluation and a possibility to participle in field -experiments.

What makes this project interesting to work on?

You will work with a highly talented team on a challenging, interesting, and practical problem using state-of-the-art deep learning methods.

What is the expected outcome?

Improving the current state-of-the-art performance and enhanced algorithm robustness.

What infrastructure, programs and tools will be used? Can they be used remotely?

Motivation and data-science tools like knowledge and experience with deep-learning. Can not be used remotely.

Is the project open source?

Part of it, and soon the rest.

What skills are necessary for this project?

Motivation and experience with deep-learning methods.

Interested candidates should be at Master, PhD or, Postdoc level.

Conformational variability analysis of small molecular structures using cryo-EM.

Dr. Tamir Bendory - Tel Aviv University, School of Electrical Engineering.

Dr. Tamir Bendory - Tel Aviv University, School of Electrical Engineering.

We work on mathematical and computational problems in data science, focusing on structural biology applications.

https://www.tau.ac.il/~bendory/

bendory@tauex.tau.ac.il

What is the project's research question?

We aim to develop methods to analyse the conformational variability of small proteins, a task which is out of reach of current technology. The project involves a variety of data science fields, including high-dimensional statistics, optimization, manifold learning, and processing of massively large data sets.

What data will be worked on?

Cryo-EM data sets available online in public repositories.

What tasks will this project involve?

Developing a new mathematical framework and implement it.

What makes this project interesting to work on?

The project has tremendous potential to revolutionize structural biology. It also involves a wide range of data science skills.

What is the expected outcome?

A new method to analyze conformational variability of small molecular structures, including an efficient and documented code that can be disseminated.

What infrastructure, programs and tools will be used? Can they be used remotely?

Infrastructure: my servers. Programs: Python/Matlab and standard cryo-EM software. Can be used remotely.

Is the project open source?

Yes

What skills are necessary for this project?

Very high mathematical and coding skills.

Interested candidates should be at Postdoc level.

Animal behaviour and animal personality.

Dr. Oren Forkosh - Cognitive and Brain Sciences Department, The Hebrew University.

Dr. Oren Forkosh - Cognitive and Brain Sciences Department, The Hebrew University.

We use AI to study animal behaviour, personalities, and emotional states in various species, including mice, cows, cats, and birds.

https://www.forkoshlab.com/

oren.forkosh@mail.huji.ac.il

What is the project's research question?

New approaches to studying animal behaviour such as identifying complex behaviours from video or position data and incorporating ideas from natural language processing to decipher animal behaviours.

What data will be worked on?

Animal position tracking data (in mice or cows) or video data (in mice or birds)

What tasks will this project involve?

Developing algorithms that involve deep learning or statistical machine learning.

What makes this project interesting to work on?

Unique data and problems (and the animals are cute)

What is the expected outcome?

New applicable algorithms to better understand animal behavior.

What infrastructure, programs and tools will be used? Can they be used remotely?

Python or Matlab as well as databases. Can be used remotely.

Is the project open source?

Partially

What skills are necessary for this project?

Python or Matlab

Interested candidates should be at Master, PhD or, Postdoc level.

Desalination Brine Reuse.

Dr. Gideon Oron - Ben Gurion University of the Negev

Dr. Gideon Oron - Ben Gurion University of the Negev

Two Experts in water and agricultural systems.

gidi@bgu.ac.il

What is the project's research question?

A contributive solution for the by-product generated during desalination.

What data will be worked on?

Amounts and qualities of the brine produced and possible solution directions.

What tasks will this project involve?

Collect data and work on modeling of reuse systems.

What makes this project interesting to work on?

Desalination is the solution for worldwide water shortage however, the brine disposal is still a problem to be considered seriously.

What is the expected outcome?

The recommend on guidelines for brine/concentrate reuse for diverse purpose.

What infrastructure, programs and tools will be used? Can they be used remotely?

Experimental systems in Sde-Boker and computing capacities. Can not be used remotely.

Is the project open source?

Yes

What skills are necessary for this project?

Background in field work and management modelling (optimization).

Interested candidates should be at Master, PhD or, Postdoc level.

Socially acceptable and fair AI

Dr. Rami Puzis - Software and Information Systems Engineering, Ben-Gurion University of the Negev

Dr. Rami Puzis - Software and Information Systems Engineering, Ben-Gurion University of the Negev

At Complex Networks Analysis Lab at Ben-Gurion University (CNALAB@BGU) we tackle research problems in diverse domains using a combination of methods from graph theory and machine learning. Complex Networks are found in cyber security, social networks, communication networks and the Internet, biological networks, financial networks, text analytics and more. Scientific programmers working the CNA Lab @ BGU develop generic software tools and libraries to analyse the structure of networks derived from the various problem domains. Graduate research students apply these tools to investigate specific problems in their domain of interest.

https://faramirp.wixsite.com/puzis

puzis@bgu.ac.il

What is the project's research question?

Many pre-trained language models exhibit bias with respect to protected classes such as gender, race, age etc. We study the fine-grained decomposition of such bias into factors well known in psychology. We try to understand to what extent do the language models exhibit the various factors of sexism, racism, etc.?

What data will be worked on?

The study population is the variety of pre-trained language models available on HuggingFace. In addition, we use datasets that exhibit specific factors of the studied biases.

What tasks will this project involve?

Fine-tuning and domain adaptation of pre-trained language models. Formulation of queries that test for specific kinds of biases and their specific factors. Analysis of results. Report/paper writing.

What makes this project interesting to work on?

The intertwining between NLP and psychology.

What is the expected outcome?

A publishable paper. A benchmark for evaluating latent constructs related to fairness and socially acceptable behavior.

What infrastructure, programs and tools will be used? Can they be used remotely?

Python, PyTorch, Transformers, Colab, GPU cluster. Can be used remotely.

Is the project open source?

Yes

What skills are necessary for this project?

Python programming. Familiarity with machine learning. Basic familiarity with NLP. Applicants should have some interest toward studying/understanding psychology.

Interested candidates should be at Master, PhD or, Postdoc level.

Multi-armed bandit algorithms for information-theoretic Channel Selection.

Dr. Nir Weinberger - Technion Israel Institute of Technology, ECE

Dr. Nir Weinberger - Technion Israel Institute of Technology, ECE

We are studying theoretical problems at the intersection of machine-learning, high dimensional statistics, and information theory.

https://sites.google.com/view/nir-weinberger/home

nirwein@technion.ac.il

What is the project's research question?

To propose and analyse exploration algorithms for finding communication channels which are information-theoretic optimal (e.g., they have maximal capacity).

What data will be worked on?

No specific data.

What tasks will this project involve?

Proposing exploration algorithms for the channel selection problem, performing a theoretical analysis, and simulating their performance.

What makes this project interesting to work on?

It involves both algorithmic questions, probabilistic analysis problems, and simulation, as well as domain knowledge in sequential decision making and information theory.

What is the expected outcome?

Algorithms for the channel selection problem, with rigorous and tight performance bounds.

What infrastructure, programs and tools will be used? Can they be used remotely?

Not applicable. Can be used remotely.

Is the project open source?

Not applicable

What skills are necessary for this project?

Mathematical maturity and basic knowledge in concentration inequalities, machine learning and information theory.

Interested candidates should be at PhD or Postdoc level.

Combining ML and Sketches for Efficient Caching

Dr. Roy Friedman - Technion, Computer Science

Dr. Roy Friedman - Technion, Computer Science

Working on stream processing and sketching algorithms, caching, and distributed systems.

https://roy.net.technion.ac.il/

roy@technion.ac.il

What is the project's research question?

Finding effective ways to combine sketches and NN to reduce the learning time as well as space and computational complexity of NN alone.

What data will be worked on?

Real-world access traces from various storage systems and Internet services.

What tasks will this project involve?

Programming the various algorithms, we come up with and running them on a cluster of GPUs to measure relative performance.

What makes this project interesting to work on?

The potential to obtain orders of magnitudes improvements in space and computational complexity.

What is the expected outcome?

Showing that we can obtain order of magnitudes improvements and publishing the results as a research paper.

What infrastructure, programs and tools will be used? Can they be used remotely?

Mostly Python/PyTorch. Hardware: A cluster of DGX-A100 machines. Can be used remotely.

Is the project open source?

Yes

What skills are necessary for this project?

Python programming, good background in ML and basic data science techniques.

Interested candidates should be at Master, PhD or, Postdoc level.

Exploring the links between outdoor exposures to green spaces and pregnancy outcomes

Dr. Keren Agay-Shay - Bar Ilan University, Azrieli Faculty of Medicine.

Dr. Keren Agay-Shay - Bar Ilan University, Azrieli Faculty of Medicine.

We are an international team of environment and health scientists studying the effects of external environmental exposures on human health and well-being. Our multidisciplinary research projects cover such scientific topics as environmental epidemiology, human biology and public health, landscape, and biodiversity studies. We apply geo-spatial data analysis and advanced statistical methods to evaluate the links between both beneficial and harmful environmental exposures and health outcomes, with the focus on adverse pregnancy outcomes (APO), maternal health during pregnancy and mental health (such as risks of stress and anxiety during pregnancy).

http://research.md.biu.ac.il/labs/keren-agay-shay/

keren.agay-shay@biu.ac.il

What is the project's research question?

To develop a novel exposure metrics that will be used to evaluate the associations between different types of residential surrounding greenness and pregnancy outcomes. Create new open source on-the-ground exposure measures using the Google Street View Imagery (SVI) and image semantic segmentation techniques.

What data will be worked on?

Main exposure data:

Participant(s) will apply geo-referenced (coordinate based) Google Street Images (SVIs) representing an on-the-ground perspective of the exposure to the external natural environment. The images captured from 2014 to 2022 on the territory of Israel will be retrieved and post-processed using image segmentation deep learning model.

Additional exposure data: satellite derived Normalized difference vegetation index (NDVI) from MODIS - Moderate Resolution Imaging Spectroradiometer (in .csv and .sav formats).

Additionally: grid shape files (.shp, .shx, .dbf formats).

What tasks will this project involve?

To quantify residential greenspace coverage (% total greenspace) from Google SVIs (selected regions in Israel) applying image segmentation model in order to estimate the percentages of each environmental class (% trees, % grass, % flowers, and % plants combined) within a 360◦ view for each given location.

To validate the accuracy of SVIs-based greenspace metrics using NDVI satellite data;

Depending on participant(s) skills set, their levels of expertise and their preferences the following technical and data analysis tasks can be performed:

creating a high-resolution regional grid using the shapefiles and/or the open street network.

extending current retrieving algorithms and adjusting Google API and street module scripts (https://github.com/robolyst/streetview) in order to obtain the locations of the images nearest to the residential addresses.

applying the pyramid scene parsing network (PSPNet) to derive greenspace metrics from the retrieved Google SVIs (pixel-level image segmentation). Each pixel within each image will be classified based on pre-trained scene parcing dataset (ADE20K);

creating geospatial raster files from segmented SVIs, which should be linked to geocoded residential addresses from pregnancy cohort data.

validation of a SVI-based metrics with the satellite-derived indicator of the quantity of vegetation on the ground (NDVI).

What makes this project interesting to work on?

The project roots in such an important topic as the benefits of public health from different types of residential landscapes. It was observed that natural residential environment including green spaces may reduce the effects of harmful exposures to rapidly increasing air temperatures and pollution in the current phase of climate change. Since for Israel there were no attempts to use machine learning techniques in greenness exposure studies, the participant(s) will contribute to the development of novel database, data processing and machine learning techniques and an overall multidisciplinary study design. The project output will allow to compliment and improve existing exposure estimates from the on-the-ground eye-level perspective. Data processing skills and machine learning techniques applied during the project can be transferred to the other fields of research. The open-source exposure metrics that will be developed will be available for future environment and health studies.

What is the expected outcome?

Database of retrieved Google Street view images for the specific study regions across Israel. Greenness exposure metrics retrieved from Google SVIs (for each given location and generalized map for the whole study region). Statistical analysis outcomes (associations of created greenness exposure metrics with adverse pregnancy outcomes).

What infrastructure, programs and tools will be used? Can they be used remotely?

Python, Jupyter-notebooks, GitLab/GitHub, Google application programming interface (API), Google Street, street view module; pyramid scene parsing network (PSPNet); pre-trained scene parcing dataset (ADE20K); of an advantage: ArcGIS, QGIS. Can be used remotely.

Is the project open source?

Yes, the geospatial exposure data (NDVI), image segmentation neural networks and pre-trained image annotation dataset as well as Google API and python modules are open source.

What skills are necessary for this project?

Scientific computation; machine learning; computer vision and image processing/analysis; programming skills (Python and/or R), data visualization, experience with GitLab/GitHub, Jupyter notebooks; Geographic information systems; of an advantage: software and databases development.

Interested candidates should be at Master, PhD or, Postdoc level.

Unsupervised learning of microbial growth curve data

Dr. Yoav Ram - School of Zoology, Faculty of Life Sciences, Tel Aviv University.

Dr. Yoav Ram - School of Zoology, Faculty of Life Sciences, Tel Aviv University.

Computational ecology and evolution.

https://www.yoavram.com

yoavram@tauex.tau.ac.il

What is the project's research question?

Microbial growth is typically assayed by measuring the density of a tube of liquid using an optical reader. The resulting measurements, called growth curves, are used to assess, and compare growth of different strains (e.g., mutants, clinical isolates, evolved strains) in various conditions (e.g., drug treatment, limited nutrients, low/high pH). These growth curves are usually studied using dynamical models, characterized by ordinary differential equations [Ram 2018]. However, these model-based approaches are limited both by model assumptions and numerical and computational challenges.

What data will be worked on?

Growth curves (OD vs time) from experiments with bacteria and fungi under different conditions.

What tasks will this project involve?

Analysis of growth curve data using classical unsupervised learning methods (e.g., PCA, k-means, hierarchical clustering [VanDerMaaten 2009]) as well as more recent methods (e.g., UMAP [McInnes 2018], variational autoencoders [Kingma 2013], autoregressive flow [Papamakarios 2017]). The student will write a final report to describe the project results, which will be assessed for its academic quality.

What makes this project interesting to work on?

Growth curves are easy to produce and automate and are therefore ubiquitous. However, they are difficult to analyse and in many cases these curves are analysed using simple methods (linear regression, area-under-the-curve), leading to loss of most of the information in the data. We will test several classical and novel machine learning methods to extract these "lost" data.

What is the expected outcome?

Completion of this project will provide the first step towards a new method for analysis of growth curves data; the participating data scientist will be invited to be involved in further development of this method and in the writing and publication of a manuscript that summarizes the new method.

What infrastructure, programs and tools will be used? Can they be used remotely?

Unsupervised learning methods (clustering, generative models, dimension reduction), including machine and deep learning. Our lab has access to a large CPU (>1000 cores) and GPU (8xA100, 24xA6000) cluster. Other students and postdocs in the lab study a variety of questions in ecology and evolution, ranging from microbes to whales to humans, applying model-based and machine learning methods to data from the lab as well as archaeological and cultural data. Can be used remotely.

Is the project open source?

Some of the data is open source.

What skills are necessary for this project?

Background in statistics/machine learning/deep learning and experience in data processing, analysis, and visualization. Knowledge of the specific methods is useful but can be covered by reading and discussions prior to the visit to our lab.

Interested candidates should be at PhD or Postdoc level.

Geometry Design for DOA Estimation in Seismic Arrays

Dr. Tirza Routtenberg - Ben-Gurion University of the Negev

Dr. Tirza Routtenberg - Ben-Gurion University of the Negev

Our research group focuses on signal processing and optimization with applications in the smart grid. We study various topics in this field, including power system data analytics, state estimation and event detection, and cyber security in the Smart Grid.

https://www.ee.bgu.ac.il/~tirzar/

tirzar@bgu.ac.il

What is the project's research question?

The research question of this data science project is: How can the direction of arrival (DOA) estimation performance of a seismic signal be optimized using a planar array sensor design, in terms of the minimum-mean-squared-periodic-error (MSPE) obtained by the commonly used maximum a-posteriori (MAP) estimator of the DOA? This research is based on comparing the MSPE of the MAP estimator with two other criteria: the Cyclic Bayesian Cramér-Rao Bound (CBCRB) and the complete Expected Log-Likelihood (ELL).

What data will be worked on?

(Comprehensive Nuclear-Test-Ban Treaty Organization) international monitoring system. It is a well-maintained and calibrated station, which ensures the high quality of the data. The data from the GEres array is continuously streamed to the IDC (International Data Centre) of the CTBTO, where it is analyzed. The data that will be used in this project has been collected by the GEres array and will be obtained from the IDC of the CTBTO.

What tasks will this project involve?

Implementing existing algorithm by Python. Testing and comparing different methods. (Potential publication, depends on the results)

What makes this project interesting to work on?

Seismic event detection and localization is an important field of study as it helps in understanding and predicting natural disasters such as earthquakes. Accurately determining the direction of arrival of seismic signals is crucial for accurate event localization.

The use of sensor array geometry in the design of the planar array for DOA estimation is an interesting and challenging problem, as it involves optimizing the performance of the estimation algorithm with respect to the sensor array geometry.

The project compares different design criteria for DOA estimation, such as the minimum-mean-squared-periodic-error (MSPE) obtained by the commonly used maximum a-posteriori (MAP) estimator, the Cyclic Bayesian Cramér-Rao Bound (CBCRB), and the complete Expected Log-Likelihood (ELL) which allows for a deeper understanding of the design problem and enables to develop better design strategies.

The use of high-quality seismic data from the GEres array, which is part of the CTBTO international monitoring system, provides an opportunity to work with real-world data and validate the proposed methodologies in a realistic scenario.

The study of the impact of different array geometry on the performance of DOA estimation could have applications in a wide range of fields, including natural disaster monitoring, explosion detection, and target localization.

What is the expected outcome?

The expected outcome of this project is to design a planar array that optimizes the direction of arrival (DOA) estimation of a narrowband signal. Thus, the expected outcome is a running algorithm and simulation platform, with possible publication (depends on the results)

What infrastructure, programs and tools will be used? Can they be used remotely?

Python. Can be used remotely.

Is the project open source?

Yes

What skills are necessary for this project?

Programming, machine learning, algorithms, signal processing (no need for all, we will design the final version of the project based on the abilities of the students)

Interested candidates should be at Master, PhD or, Postdoc level.

Deep Learning Approach for Countrywide High-Resolution Air Pollution Modelling.

Dr. Avitay Geltman - Technion - Israel Institute of Technology

Dr. Avitay Geltman - Technion - Israel Institute of Technology

Devising machine & deep learning methods and mathematical models for better understanding built and natural complex environments.

https://fishbain.net.technion.ac.il/

avitay@campus.technion.ac.il

What is the project's research question?

Predicting and modelling country-scale air pollution levels using advanced neural network architectures (Transformers/diffusion models/GANs, etc., yet to be decided).

What data will be worked on?

Time-series tabular data, containing key variables such as: Monitored pollutant concentrations, transportation and industry emissions, meteorological variables, predictions of a deployed Physico-chemical model.

What tasks will this project involve?

Data pre-processing, implementation of feature selection methods, development of neural network architectures, analyzing results and deriving conclusions.

What makes this project interesting to work on?

The project involves modern promising deep learning methods based on high-quality real-world data. A successful development would be eligible for deployment by the Israeli Ministry of Environmental Protection.

What is the expected outcome?

A publication + a potential deployment of the model in Israel.

What infrastructure, programs and tools will be used? Can they be used remotely?

Python, PyTorch (/ PyTorch Lightning), etc. Can be used remotely.

Is the project open source?

Partially

What skills are necessary for this project?

Background in machine & deep learning, with research experience in the field.

Interested candidates should be at Master, PhD or, Postdoc level.

Medical procedure length of stay and how it affects schedule variability using robust optimization methods versus current practices.

Dr. Noam Goldberg - Bar-Ilan University, Department of Management

Dr. Noam Goldberg - Bar-Ilan University, Department of Management

Data driven optimization, robust optimization.

https://management.biu.ac.il/en/Noam.Goldberg

noam.goldberg@biu.ac.il

What is the project's research question?

Can utilization and schedule variability be improved using robust scheduling and bin packing techniques.

What data will be worked on?

Anonymized hospital visits and appointment data.

What tasks will this project involve?

Analysis of length of stay anonymized data for different hospital departments and medical procedures.

What makes this project interesting to work on?

Working with data analysis and visualization tools to get insights into real life processes. Interesting optimization algorithms that can lead to better decision making. Working on important healthcare applications.

What is the expected outcome?

Implementation of robust scheduling and bin packing algorithms. Analysis and characterization of LOS distribution.

What infrastructure, programs and tools will be used? Can they be used remotely?

R, Python, Julia. Can mostly be used remotely.

Is the project open source?

Partially

What skills are necessary for this project?

Programming skills in Python and/or Julia (preferred)

Interested candidates should be at PhD level.

Three-dimensional scene reconstruction using echolocation echoes

Dr. Igal Bilik - Ben Gurion University. School of Electrical and Computer

Dr. Igal Bilik - Ben Gurion University. School of Electrical and Computer

Statistical and machine learning-based signal processing for sensors and sensor arrays enabling autonomous platforms operation. Smart Sensing Lab research focuses on signal processing for radar, visual and acoustic sensors data.

https://danielbilik2003.wixsite.com/igalbilik

bilik@bgu.ac.il

What is the project's research question?

Detection, Localization, and classification of human operating in adjacency with robotic arm.

What data will be worked on?

Vision and radar recordings of human and robotic arm operating in laboratory environment.

What tasks will this project involve?

Development of Deep learning-based approach for detection, localization, and classification of human and robotic arm.

What makes this project interesting to work on?

Ability to localize and classify between human and robotic arm is critical to introduction of the consumer cobots into human living environment. Reliable algorithms are still missing, and this project can provide a basis for the new era of robot-human interaction. The project can have significant academic and industrial implications.

What is the expected outcome?

Journal article summarizes derived algorithms and presenting results.

What infrastructure, programs and tools will be used? Can they be used remotely?

Computer with GPUs , Python with DLL libraries. Can be used remotely.

Is the project open source?

No

What skills are necessary for this project?

Statistical Signal Processing, Computer vision, Deep learning.

Interested candidates should be at Master, PhD or, Postdoc level.