Ameen Eetemadi

Title: Bioinformatics Expert
Email: eetemadi@ucdavis.edu
Research Area: Biomedical Informatics and Machine Learning

UC Davis Genome Center and
Department of Computer Science,
Kemper Hall 2063,
University of California, Davis

Selected Publications

- **Eetemadi, A.** and Tagkopoulos, I., 2021. Algorithmic Lifestyle Optimization: rapid discovery of effective lifestyle interventions in individuals using group testing. *Under Review*
- **Eetemadi, A.** and Tagkopoulos, I., 2021. Methane and fatty acid metabolism pathways are predictive of Low-FODMAP diet efficacy for patients with Irritable Bowel Syndrome. *Clinical Nutrition*
- Wang X, Rai N.,Pereira B., **Eetemadi, A.** and Tagkopoulos, I., 2020. Accelerated knowledge discovery from omics data by optimal experimental design. *Nature Communications*
- **Eetemadi, A.**, Rai N., Pereira B., Kim M., Schmitz H. and Tagkopoulos, I., 2020, The Computational Diet: A Review of Computational Methods Across Diet, Microbiome, and Health. *Frontiers in Microbiology*
- **Eetemadi, A.** and Tagkopoulos, I., 2019. Genetic Neural Networks: an artificial neural network architecture for capturing gene expression relationships. *Bioinformatics*.
- Kim, M.*, **Eetemadi, A.*** and Tagkopoulos, I., 2017. DeepPep: Deep proteome inference from peptide profiles. *PLoS computational biology*, 13(9), p.e1005661. (***contributed equally**)
- **Eetemadi, A.**, 2012. Medical data analysis method for epilepsy, Master's dissertation, *Wayne State University*.
- **Eetemadi, A.**, Siadat, M.R., Soltanian-Zadeh, H., Fotouhi, F. and Elisevich, K., 2007, "Content-Based Support Environment (C-BASE): Data Preparation and Similarity Measurement.",*Proceedings of the Seventh IEEE International Conference on Data Mining (ICDM'07)*, pp. 145-150, Omaha, NE, USA, October 28-31,

Education

- **University of California**, Davis, CA (graduated 2021)
  Ph.D. in Computer Science (Designated Emphasis in Biotechnology)
- **Wayne State University**, Detroit MI (graduated 2012)
  M.Sc in Computer Science (Midical Data Mining)
- **Sharif University of Technology**, Tehran, Iran (graduated 2005)
  B.Sc in Computer Engineering (Software)

Work experience

- **University of California**, Davis, CA (2014 - 2021)
  Graduate Research Assistant, Department of Computer Science and Genome Center
- **Microsoft**, Redmond, WA (2008 - 2014)
  Software Development Engineer, Microsoft Office
- **Henry Ford Health Systems**, Detroit, MI (2005 - 2008)
  Graduate Research Assistant, Health Informatics

Teaching experience

- **University of California**, Davis, CA (2014 - 2020)
  Instructor: ECS 124 Bioinformatics
  Teaching Assistant: ECS 171 Machine Learning, ECS 124 Bioinformatics, ECS 120 Theory of Computations, ECS 36C Data Structures and Algorithms, ECS 30 Programming&Prob Solving.

Skills

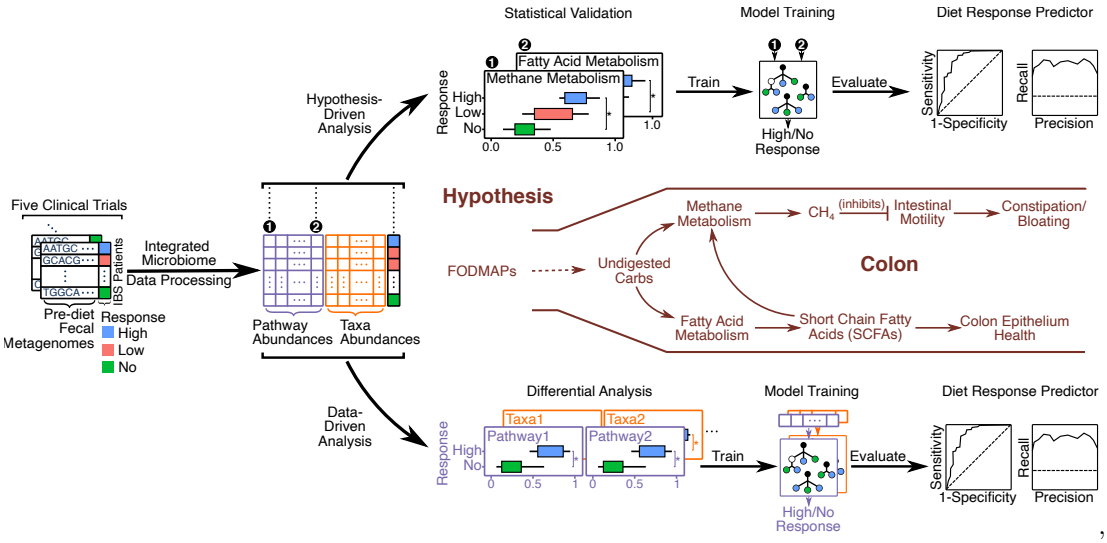
- **Software Technologies**
  Sequence analysis (DNA, RNA and Metagenomics), Supervised and Unsupervised Machine Learning (Torch7, TensorFlow and scikit-learn), TCP/IP and RESTful web services.
- **Programming Technologies**
  Python, R, MATLAB, C++, C#, Java, HTML JavaScript, AngularJS, Node.js, Lua, ASP.net, PHP, HPC (Slurm, TORQUE), Gurobi Optimizer
- **Database Systems**
  MongoDB, PostgreSQL, Oracle (+PL/SQL), MSSQL

Selected Publication Abstracts

Methane and fatty acid metabolism pathways are predictive of Low-FODMAP diet efficacy for patients with Irritable Bowel Syndrome

Ameen Eetemadi and Ilias Tagkopoulos.
Clinical Nutrition (2021).
https://doi.org/10.1016/j.clnu.2020.12.041.

Abstract: *Objective.* Identification of microbiota-based biomarkers as predictors of low-FODMAP diet response and design of a diet recommendation strategy for IBS patients. *Design.* We created a compendium of gut microbiome and disease severity data before and after a low-FODMAP diet treatment from published studies followed by unified data processing, statistical analysis and predictive modeling. We employed data-driven methods that solely rely on the compendium data, as well as hypothesis-driven methods that focus on methane and short chain fatty acid (SCFA) metabolism pathways that were implicated in the disease etiology. *Results.* The patient's response to a low-FODMAP diet was predictable using their pre-diet fecal samples with F1 accuracy scores of 0.750 and 0.875 achieved through data-driven and hypothesis-driven predictors, respectively. The fecal microbiome of patients with high response had higher abundance of methane and SCFA metabolism pathways compared to patients with no response (p-values $< 6 \times 10^{-3}$). The genera *Ruminococcus 1*, *Ruminococcaceae UCG-002* and *Anaerostipes* can be used as predictive biomarkers of diet response. Furthermore, the low-FODMAP diet followers were identifiable given their microbiome data (F1-score of 0.656). *Conclusion.* Our integrated data analysis results argue that there are two types of patients, those with high colonic methane and SCFA production, who will respond well on a low-FODMAP diet, and all others, who would benefit a dietary supplementation containing butyrate and propionate, as well as probiotics with SCFA-producing bacteria, such as *Lactobacillus*. This work demonstrates how data integration can lead to novel discoveries and paves the way towards personalized diet recommendations for IBS. The source code for reproducing the article results is available at https://github.com/IBPA/FODMAPsAndGutMicrobiome.

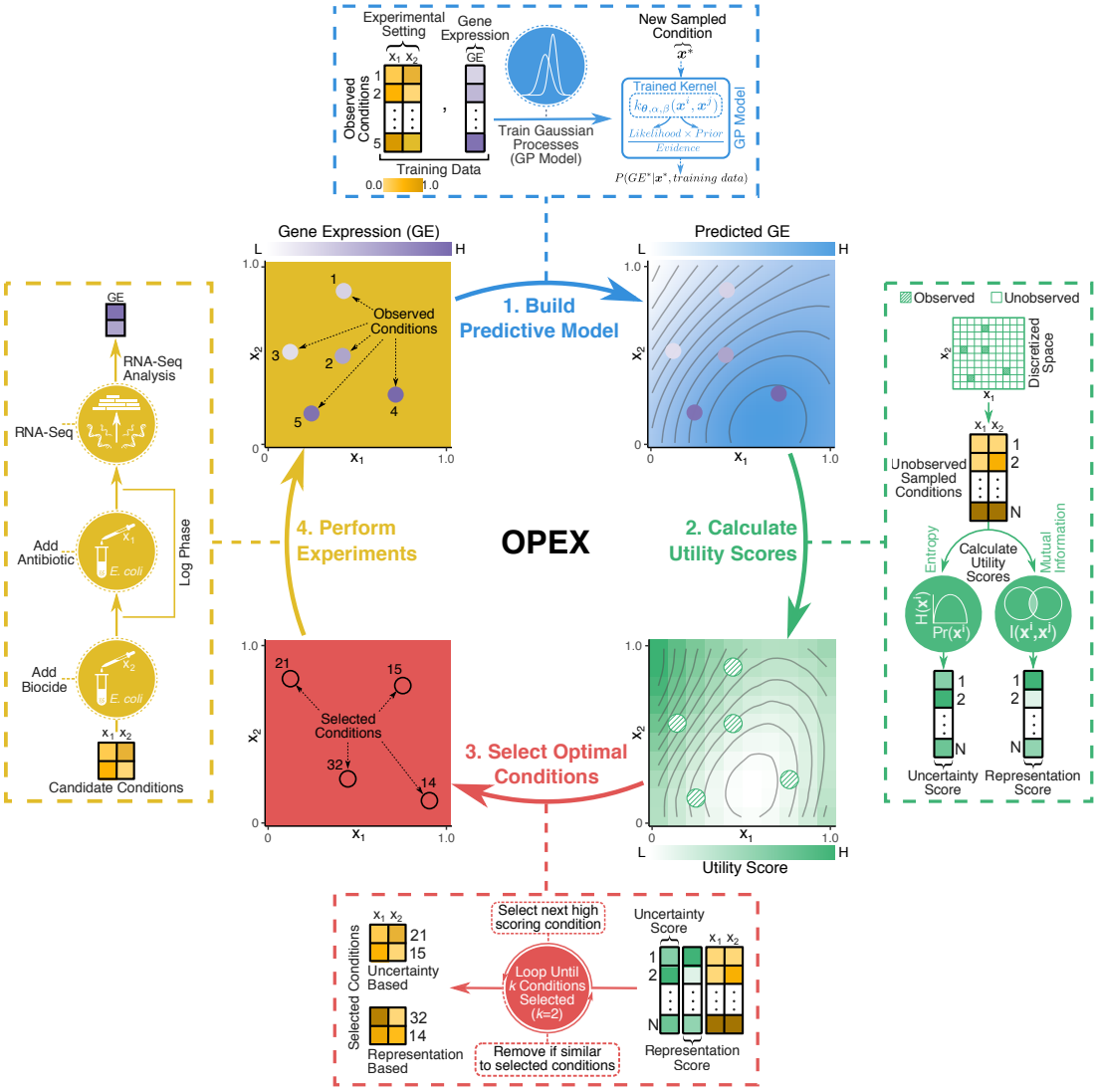


.

Accelerated knowledge discovery from omics data by optimal experimental design

Xiaokang Wang, Navneet Rai, Beatriz Merchel Piovesan Pereira, **Ameen Eetemadi**, and Ilias Tagkopoulos.
Nature Communications 11.1 (Oct. 2020), p. 5026.
https://doi.org/10.1038/s41467-020-18785-y.

Abstract: How to design experiments that accelerate knowledge discovery on complex biological landscapes remains a tantalizing question. We present an optimal experimental design method (OPEX) to identify informative omics experiments using machine learning models for both experimental space exploration and model training. OPEX-guided exploration of *Escherichia coli*'s populations exposed to biocide and antibiotic combinations lead to more accurate predictive models of gene expression with 44% less data. Analysis of the proposed experiments shows that broad exploration of the experimental space followed by fine-tuning emerges as the optimal strategy. Additionally, analysis of the experimental data reveals 29 cases of cross-stress protection and 4 cases of cross-stress vulnerability. Further validation reveals the central role of chaperones, stress response proteins and transport pumps in cross-stress exposure. This work demonstrates how active learning can be used to guide omics data collection for training predictive models, making evidence-driven decisions and accelerating knowledge discovery in life sciences.

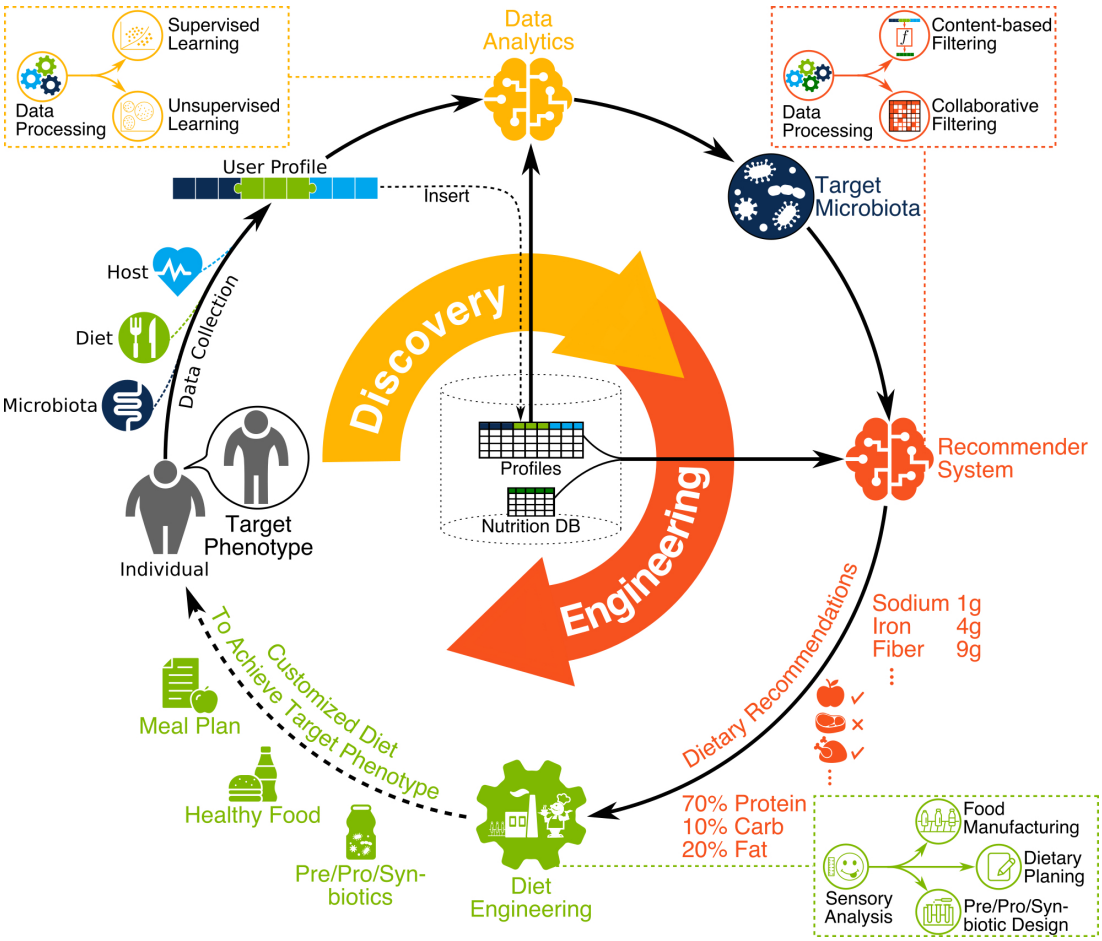


.

The Computational Diet: A Review of Computational Methods Across Diet, Microbiome, and Health

Ameen Eetemadi, Navneet Rai, Beatriz Merchel Piovesan Pereira, Minseung Kim, Harold Schmitz, and Ilias Tagkopoulos.
Frontiers in Microbiology 11 (2020), p. 393.
https://www.frontiersin.org/article/10.3389/fmicb.2020.00393.

Abstract: Food and human health are inextricably linked. As such, revolutionary impacts on health have been derived from advances in the production and distribution of food relating to food safety and fortification with micronutrients. During the past two decades, it has become apparent that the human microbiome has the potential to modulate health, including in ways that may be related to diet and the composition of specific foods. Despite the excitement and potential surrounding this area, the complexity of the gut microbiome, the chemical composition of food, and their interplay in situ remains a daunting task to fully understand. However, recent advances in high-throughput sequencing, metabolomics profiling, compositional analysis of food, and the emergence of electronic health records provide new sources of data that can contribute to addressing this challenge. Computational science will play an essential role in this effort as it will provide the foundation to integrate these data layers and derive insights capable of revealing and understanding the complex interactions between diet, gut microbiome, and health. Here, we review the current knowledge on diet-health-gut microbiota, relevant data sources, bioinformatics tools, machine learning capabilities, as well as the intellectual property and legislative regulatory landscape. We provide guidance on employing machine learning and data analytics, identify gaps in current methods, and describe new scenarios to be unlocked in the next few years in the context of current knowledge.

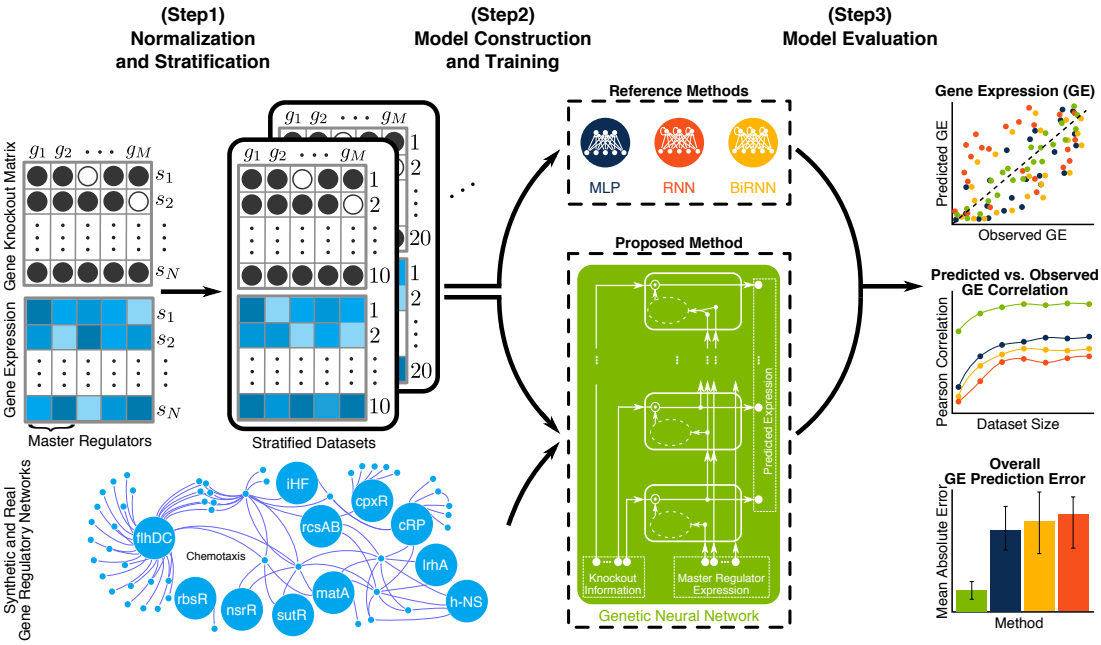


.

Genetic Neural Networks: an artificial neural network architecture for capturing gene expression relationships

Ameen Eetemadi and Ilias Tagkopoulos.
Bioinformatics 35.13 (2018), pp. 2226–2234.
https://doi.org/10.1093/bioinformatics/bty945.

Abstract: Gene expression prediction is one of the grand challenges in computational biology. The availability of transcriptomics data combined with recent advances in artificial neural networks provide an unprecedented opportunity to create predictive models of gene expression with far reaching applications. We present the Genetic Neural Network (GNN), an artificial neural network for predicting genome-wide gene expression given gene knockouts and master regulator perturbations. In its core, the GNN maps existing gene regulatory information in its architecture and it uses cell nodes that have been specifically designed to capture the dependencies and non-linear dynamics that exist in gene networks. These two key features make the GNN architecture capable to capture complex relationships without the need of large training datasets. As a result, GNNs were 40% more accurate on average than competing architectures (MLP, RNN, BiRNN) when compared on hundreds of curated and inferred transcription modules. Our results argue that GNNs can become the architecture of choice when building predictors of gene expression from exponentially growing corpus of genome-wide transcriptomics data. The source code of GNN is available at https://github.com/IBPA/GNN.



.

DeepPep: Deep proteome inference from peptide profiles

Minseung Kim, **Ameen Eetemadi**, and Ilias Tagkopoulos.
PLOS Computational Biology 13.9 (2017), pp. 1–17.
https://doi.org/10.1371/journal.pcbi.1005661.

Abstract: Protein inference, the identification of the protein set that is the origin of a given peptide profile, is a fundamental challenge in proteomics. We present DeepPep, a deep-convolutional neural network framework that predicts the protein set from a proteomics mixture, given the sequence universe of possible proteins and a target peptide profile. In its core, DeepPep quantifies the change in probabilistic score of peptide-spectrum matches in the presence or absence of a specific protein, hence selecting as candidate proteins with the largest impact to the peptide profile. Application of the method across datasets argues for its competitive predictive ability (AUC of 0.80±0.18, AUPR of 0.84±0.28) in inferring proteins without need of peptide detectability on which the most competitive methods rely. We find that the convolutional neural network architecture outperforms the traditional artificial neural network architectures without convolution layers in protein inference. We expect that similar deep learning architectures that allow learning nonlinear patterns can be further extended to problems in metagenome profiling and cell type inference. The source code of DeepPep and the benchmark datasets used in this study are available at https://deeppep.github.io/DeepPep/.