EXPLAINABLE-AI IN BIOTECHNOLOGY & PHARMACEUTICAL

Overview

When it comes to drugs and medical treatments there is no 'one-size fits all'. Patients vary greatly in their needs and responses. A treatment that is life-saving for one person might be ineffective or even harmful to another. This realization has led to the revolution of personalized medicine and drug design.

It is helpful here that there is a certain degree of commonality among people. Much can be learned from partitioning the overall population of patients into subpopulations that share certain common features and attributes. However, identifying well-defined subpopulations remains to be a challenging endeavor.

While there is nowadays no shortage in data that can be used to minutely characterize individual patients and the symptoms they exhibit, these detailed characterizations lead to vast and unwieldy feature spaces where patient subpopulations are rarely homogeneous and often difficult to separate.

Furthermore, many of the features and attributes may not be important for a particular task but this is difficult to assess beforehand. Practitioners are often riddled with the problem of selecting the right features for a specific analysis problem. Unable to cope, feature selection reduces to a guessing game.

There are many application areas in the BioTech/Pharma sector where these critical challenges occur:

- **Rare diseases:** select the most promising treatment for a given patient
- **Treatment prognosis:** match drug interventions with individual patients
- **Drug testing and validation:** select the most appropriate candidates for a clinical trial
- **Drug re-purposing:** find new associations of disease progression patterns for a given drug

Typical Challenges:

In any of these applications, as well as others, knowing a data point's membership in a concrete and statistically robust subpopulation is critically important. It is because determining a point's membership with high confidence drastically reduces risk and uncertainty in the decision's outcome.

Moreover, many times it is not only important to know whether a particular data item will yield a certain outcome in a given response variable. Rather, it can also be important whether it allows certain predictions with regards to the variable's growth or decline. For example:

- Will a patient experience a drop in blood pressure as the level of medication is increased

- Will a given individual have adverse reactions to an administered drug and if so, how much

- Will a selected trial candidate demonstrate a significant response to the tested drug

- What is the margin of safety of a certain dose in relation to its benefit

These kinds of predictions are commonly assessed by statistically correlating the target variable with one or more attributes. However, correlation can be sporadic and confined to a certain subpopulation only. It may even vanish and remain undetected when aggregated over loosely defined subpopulations.

Identifying robust and well defined subpopulations in large datasets is hence mission-critical. However, the aforementioned "data fog" in which these subpopulations are typically situated poses tremendous challenges. In other words, large data is a blessing and a curse at the same time.

Deep neural networks, random forests, etc. circumvent this problem in an elegant manner but this comes at the price of low explainability and subsequent poorly informed decision making. This uncertainty is a fateful adversary in the high stakes effort of personalized medicine and drug design.
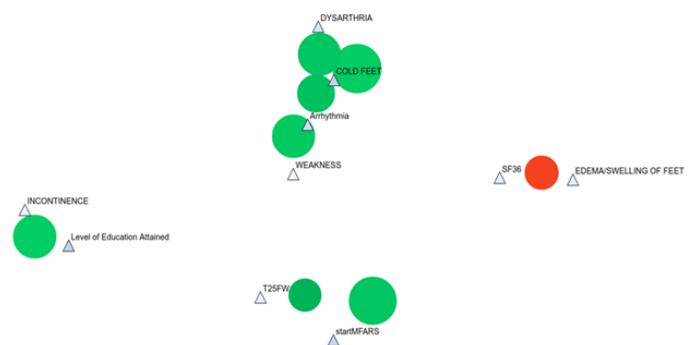
Contextual Pattern Visualization
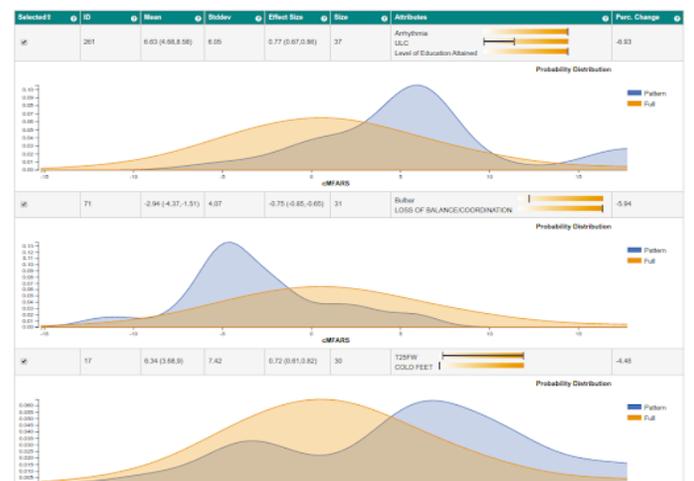
Salient Pattern Mining Output

How Explainable-AI can help:

Akai Kaeru's advanced pattern mining and discovery engine operates at the root of the problem. It directly decomposes the high-dimensional input data into a reliable set of independent data patterns -- the subpopulations -- using a sophisticated yet efficient suite of proprietary algorithms.

Each pattern consists of data items that behave similarly in terms of a given target variable and which are succinctly defined by just a small set of attributes. The succinct description of each pattern makes them easy to understand and appeals to the user's domain knowledge – a hallmark of explainable AI.

Our pattern mining engine delivers its products to a novel fully interactive visual interface. It alerts the analyst to a set of concrete and concise regions with unusually high or low target variable values and correlations. The patterns are explained by visualizing them in the context of their relevant attributes.

Also part of our software is a causal inference engine which can establish true causal relations among the patterns. It provides directional relations such as "smoking causes cancer, but not vice versa". Our visualization interface conveys the intricate web of causal relationships directly on the pattern layout.
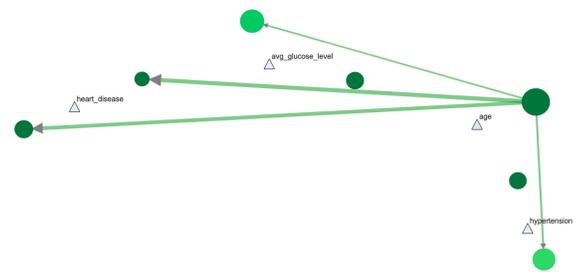
Via these capabilities, Akai Kaeru's software can help analysts and decision makers turn troves of data into actionable insight they can trust, justify and explain to others. There are no black magic boxes
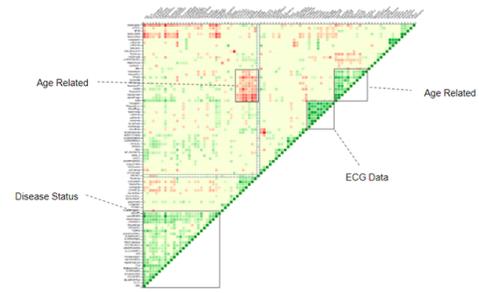
Correlation Heat Map

Probability Bar Chart

Causal Relationship Display

About Us

Zeblok Computational Inc. provides the first turnkey Supercomputing-as-a-Service to make AI and data-science ubiquitously available to enterprises large and small, all the way to the edge data centers. We make it simple and cost effective to build AI-applications by combining smart-sourced surplus infrastructure capacity with a curated library of AI algorithms sourced from top experts in the world.

Zeblok's AI-Workstation builds upon and virtualizes the ubiquitous Jupyter Notebook technology, supporting all popular AI/ML frameworks as well as popular languages like R, Python and Scala, enabling data scientists to begin creating customized applications in a pre-configured environment set up in minutes at the click of a button. Under the hood is a powerful multi-cloud, multi-class orchestration, scheduling and optimization engine to accommodate workloads ranging from a single GPU to hundreds.

The **Zeblok AlgoStore** is a library of curated original algorithms, which can be used to train models, and are ready to be integrated into enterprise business processes, with Zeblok engineers able to aid in customization.

Akai Kaeru, LLC creates AI-powered software that helps data scientists solve stubborn problems in real-world applications. Leveraging extensive research expertise in interactive high-dimensional data visualization, its two co-founders transformed acclaimed research products into the practical Salient Pattern Miner, Visual Causal Analyst, and Data Context Map that are at the heart of the Explainable AI data analytics software suite.

Zeblok and **Akai Kearu** have come together to offer the Explainable-AI algorithm as part of Zeblok's ecosystem.

More information on Akai Kearu at https://akaikaeru.com/