www.trendalyze.com

White Paper

Scientific Approach of Visual Motif Discovery

MOTIF DISCOVERY OFFERS TREMENDOUS OPPORTUNINITES TO IMPROVE RESEARCH AND OUTCOMES MANAGEMENT IN HEALTHCARE

KEY POINTS:

Visual Motif Discovery uses a heuristic search approach that everyone understands.

With visual motif discovery a clinician can see a pattern once and search patients for occurrences elsewhere.

Trendalyze uses search to detect patterns, anomalies and perform analysis and monitoring against real time data from thousands or millions of patients.

We need scalability to support an entire health system such as the NHS in England or the VA in the United States.

With Trendalyze we can have the human in the loop and avoid all the time, cost and complexity of the computer programming.

TRENDALYZE

Scientific Approach of Visual Motif Discovery

Visual Motif Discovery uses a heuristic search approach that everyone understands. We all use Google and Bing to search for words and phrases in documents and web pages. The analogy for wearables and medical sensors is searching time series data streams which require a visual motif discovery and search. Clinicians and academic researchers are very familiar with charts showing time series (patterns) of vital signs. With visual motif discovery a clinician can see a pattern once and search patients for occurrences elsewhere. Trendalyze uses search to detect patterns, anomalies and perform analysis and monitoring against real time data from thousands or millions of patients. We call these repeated trends - motifs. The clinician can use an iPhone, iPad or Web browser to browse through the charts to interactively spot something of interest, a spike, a repetition, a increasing trend - the motif. Next, the clinician defines a search domain - where to look for the motifs and executes a search which can be against individual patient, a patient list or all patients. The search results immediately show insights which can be filtered, refined and correlated. Such search results and the motifs can be shared with other clinicians for collaboration and validation. Patients can be monitored in real-time.

So what is the motif search? It gives ranked order of patient matches based on some predefined distances or geometric shapes. We have implemented several distance and pattern search algorithms to be used depending on the clinical use cases. Some of them are fit for closeness of actual values, some are better for shape similarity. Motifs and time series domains can be multi measure, e.g. systolic and diastolic blood pressure, or min and max temperatures or derived measures such as the NHS National Early Warning Score (NEWS) or Mean Blood Pressure (MBP). Time correlated motif searches allow us to match different events over time, for example a sudden reduction in blood pressure, an increasing temperature and a decline in heart rate.

A key aspect is how the search is executed or how do we go over the search domain. We need scalability to support an entire health system such as the NHS in England or the VA in the United States. Firstly, the algorithms are designed so that the number of iterations are linearly dependent with the size of the search domain. We can actually do better due to vector pipeline processing, in-memory columnar storage and sort merge join processing. Secondly, the use the Apache Spark computational engine, whose strength is to execute algorithms in parallel on cloud computing platforms gives a huge performance boost. One of our insights was that searching for motifs is easier than statistics. Clinical outcomes are visually described by motifs in the time series, so the way to understand their causes is through searching for correlated trends, anomaly and geometric shapes. Furthermore, statistical predictions generate a lot of false positives. Trendalyze has use case for remote patient monitoring for congestive heart failure (CHF) patients where false positives are almost completely missing.

Comparison of Motif Search to Machine Learning/Artificial Intelligence

Machine learning algorithms are built with two assumptions: that there is a data model and that a training data set exists. These are always problematic. The model is idealised version of reality and might be either incorrect or restrictive due to sampling. The training data set can be too small to make adequate predictions or even non existing. These algorithms are usually tailor made for a specific problem and a little tweak in the desired results needs a large development effort and time.

Where is Trendalyze strength? There is the ability of the human eyes (the brain of the subject matter expert) to spot a trend or anomaly in data using interactive charts/visualizations something clinicians do every day using charts. But it is hard for the human to detect repeating patterns (motifs) in big data sets (for example thousands of patients or millions of patients) that are key to understanding the root causes of events. Trendalyze unifies of the above two. A clinician can try different scenarios until finding events and apply motif search for millions of other patients. Also, Trendalyze is not domain specific and experts in every domain can perform the same process to make deductions based on their specific data. For example, the same heuristic search approach and visual tools can be used for detecting hypertension, chronic pain flare-ups and CHF conditions.

So how does Trendalyze compare with machine learning approaches? The difference is the role of clinical domain subject matter expert (such as consultant cardiologist or academic researcher) in the process and how the visual tools coupled with search make the human eye many more times scalable. The clinician can browse through structured data using charts to spot a known phenomenon or make a conjecture based on relevant experience. The search will help find other examples in the first case and confirm or deny ideas in the second. And then the process can iterate until all secrets of the data are revealed. Then Trendalyze can monitor patients in real time. So there is no model, just the human insight from the clinician subject matter expert, which makes it dramatically simpler and more effective. Trendalyze will distinguish the effective analysts and give invaluable insights to help better manage clinical outcomes.

By contrast in the classic Machine Learning approach models have to be developed by computer programmers iteratively tested and validated. This adds cost and complexity and is no match for the power of the knowledge of the clinical subject matter expert or the academic researcher. With Trendalyze we can have the human in the loop and avoid all the time, cost and complexity of the computer programming.

The other issue was the training data. Machine learning needs to learn from somewhere, typically a large set of preexisting data. An expert in a particular clinical domain can always do better than a such a limited model. Trendalyze can have one patient and start making predictions on the events to follow. And the solution can scale to the requirements of an entire health network.

Example: A clinician looks at vital sign charts of 1000 random sample of patients and gets patterns of interest. Spending an average of 10 seconds on each patient would make it 3 hours for initial analysis. Trendalyze can then search the universe of patients (say one million patients) for similar patterns in a few seconds. The search can be iterative process to refine searches based on new patterns, trends and anomaly that are discovered during the search process. It would take 4 months for a human to manually view one million patient charts (working 24 hours a day, 7 days a week). This is obviously not practical and leads to the approaches of statistical sampling which tends to exclude real insight possible through analyzing each individual patient. The Trendalyze approach reduces this to 3 hours initial analysis time and then repeated search and monitors in seconds against all patients. This can decrease operational costs, improve quality of care, improve clinical outcomes and prevent mortalities.

Augmenting Motif Discovery and Search with Deep Learning

Recently Trendalyze has integrated some leading edge deep learning technology (TensorFlow from Google) to augment the human visual discovery using neural networks. This allows the training of networks based on the data streams and the motifs selected by the users and the searches. The network can be trained to predict the likely motifs in future which can be visualised and leveraged by the clinical subject matter expert in the discovery process. We are currently testing this approach with a University College London (UCL) image guided surgery project for robotic assisted prostatectomy procedure. However, the same application and data analytics can be applied in other domains for example general robotic surgery and motion patterns in urology, transcatheter aortic valve replacement (TAVI) procedure, NEWS in telehealth remote CHF patient monitoring and real-time vital signs in acute hospital OR/ICU settings e.g patients under anesthesia.

More details on the ERDF funded project at www.trendalyze.com