Apexon

WHITE PAPER

APPLYING MACHINE LEARNING IN HEALTHCARE

EXTRACTING INSIGHTS AND VALUE OUT OF LARGE, COMPLEX DATA SETS WITHIN THE HEALTHCARE DOMAIN

apexon.com

INTRODUCTION

TAKING KNOWLEDGE DISCOVERY TO THE NEXT LEVEL

The concept of “Big Data” has worn out its welcome in many enterprises these days as users grow weary of the high costs and complexity involved with structuring and storing huge volumes of data in clustered environments.

It covers aspects of the AWS as a development platform, AWS cloud products and solutions, building cloud native applications, advantages of AWS, support, pipelines, automation, testing tools and techniques.

KNOWLEDGE DISCOVERY

THE RISE OF MACHINE LEARNING

It has value where it can provide timely high-value and actionable insights, or influence workflows in a highly automated fashion, but those opportunities are very limited for the majority of organizations.

But the need for knowledge discovery has not disappeared. In Big Data's place, we've seen the rise of machine learning. This is especially true in healthcare, where machine learning is being applied very effectively; e.g.:

- Accurately identifying noise vs. real signals in information using cognitive learning and prediction.
- Finding hidden patterns which are beyond human assessment capacity.
- Augmenting human capabilities and filling the gap where humans may be limited related to information overload, analytical accuracy, and empirical judgement.
- Supporting and accelerating innovations in speech recognition, advanced search, and image processing.
- Processing repetitive data management tasks more quickly, efficiently, and accurately.

KNOWLEDGE DISCOVERY

A GUIDED TOUR OF MACHINE-LEARNING ALGORITHMS IN HEALTHCARE

There are many different machine-learning algorithms – distinguished by categories, data types, business problems and objectives. And there is a substantial amount of mathematical and statistical underpinning for each.

But at its core, machine learning is solving human problems and these algorithms fit into the business world quite naturally. Below we look at the different types of machine learning algorithms and how they are being applied to address a number of very practical and valuable use cases in healthcare.

“To be useful, data must be analyzed, interpreted, and acted on… [and] attention has to shift to new statistical tools from the field of machine learning that will be critical for anyone practicing medicine in the 21st century.”

Obermeyer, Z. and Emanuel, E.J. (2016, September 29). “Predicting the Future – Big Data, Machine Learning, and Clinical Medicine.” New England Journal of Medicine, Vol. 375, p. 1216.

CLASSIFICATION

WHICH CLASS DOES DATA BELONG TO?

Example: Is he really a heart patient?

There are several different types of Classification algorithms:

LOGISTIC REGRESSION

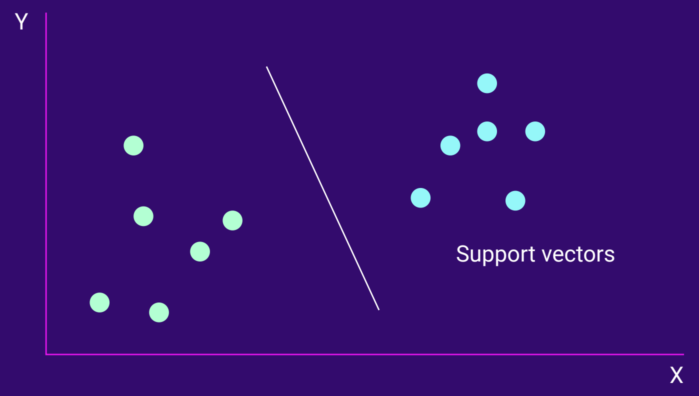
A statistical method for analyzing a dataset in which there are one or more independent variables that determine an outcome – measured with a dichotomous variable (only two possible outcomes). In logistic regression, the dependent variable is binary or dichotomous, i.e. it only contains data coded as 1 (TRUE, success, pregnant, etc.) or 0 (FALSE, failure, non-pregnant, etc.).

DECISION TREES, RANDOM FOREST, BOOSTED TREES

An entire family of algorithms, all based on the idea of creating a tree of decisions about features that lead to a specific classification. Decision trees build classification or regression models in the form of a tree structure — breaking down a dataset into smaller and smaller subsets while developing an associated decision tree incrementally at the same time. The final result is a tree with decision nodes and leaf nodes. A decision node has two or more branches. Leaf nodes represents a classification or decision. The topmost decision node in a tree corresponds to the best predictor called a root node. Decision trees can handle both categorical and numerical data. There are also plenty of ways in which decision trees work for additional analysis like splitting, pruning, etc.

SUPPORT VECTOR MACHINE (SVM)

A supervised machine learning algorithm that can be employed for both classification and regression purposes. SVMs are most commonly used in classification problems – the idea of finding a hyperplane that best divides a dataset into two classes, as shown in the image below.



DEEP LEARNING

Deep learning (and neural networks in general) can take raw data as input and produce a class (or a vector of probabilities for many classes) as output. All neural network models consist of multiple layers of "neurons." Each neuron in a layer receives the outputs of neurons in previous layers, combines these inputs, and uses a threshold to determine whether to output a value closer to 0 or closer to 1 for processing by the next layer. Deep-learning algorithms are unique in their ability to automatically generate features of interest in input data, as long as they are provided with a sufficient number of training examples (usually in the millions. although this is changing with new advances in the area of Transfer Learning). Deep learning is already used extensively in image, video, and audio understanding in healthcare, and it will eventually become more common in other classification problems in healthcare as larger sets of labeled training data become available for healthcare applications.

COMMON USE CASES

Classifiers are the most commonly used machine learning algorithms in all analytics applications, including healthcare.

Some specific use cases include:

- Suggesting possible patient diagnoses
- Identifying patients with high readmission risk
- Automatically alerting care providers early in the development of sepsis
- Defining thresholds for "abnormal" lab results
- Automatically differentiating between clinical and administrative documents
- Recommending the most effective wellness or disease management intervention for a patient

MEMORY-BASED LEARNING

DOES THIS DATA SOUND LIKE A PAST PATTERN?

Example: Do we have past patients that resemble this patient?

ASSOCIATIVE MEMORY

Associative memory systems compare incoming data with past data to create a pattern resemblance and to understand what the incoming data is all about. Comparisons can be based on any number of data attributes and have a lot of flexibility in that aspect. This is especially useful in healthcare where we typically have a large number of attributes. There are many measurements for each curation and each test generally has numerous data points.

COMMON USE CASES

- Insurance fraud
- Avoiding readmission costs
- Patient cohort creation for reference; crucial for identifying the buildup of a specific disease like patterns of patients moving from diabetes 1 to diabetes 2 and related eyesight loss

FORECASTING

HOW MUCH THIS TIME SERIES CHANGE IN NEXT TIME PERIOD?

Example: How likely this diabetes patient will be able to progress in future?

LINEAR REGRESSION

Linear regression attempts to model the relationship between two variables by fitting a linear equation to observed data. One variable is considered to be an explanatory variable, and the other is considered to be a dependent variable. For example, a modeler might want to relate the weights of individuals to their heights using a linear regression model.

NEURAL NETWORKS

A neural network takes past values as inputs and produces the predicted next value as output. These can be as simple as a Multi-Layer Perceptron (MLP), or as complex as a recurrent deep-learning model (e.g., Long Short-Term Memory, LSTM). All neural network models consist of multiple layers of "neurons." Each neuron in a layer receives the outputs of neurons in previous layers, combines these inputs, and uses a threshold to determine whether to output a value closer to 0 or closer to 1 for processing by the next layer.

COMMON USE CASES

- Risk prediction for individuals
- Time-sensitive disease detection
- Chase list for chronic disease management: predictive future values of key variables like haemoglobin or blood pressure to predict potential candidates for chronic disease management

PROBABILITY ESTIMATION

MOST LIKELY INTERPRETATION OF DATA

Example: What's the most probable diagnosis decision given various signs of patient?

LOGISTIC REGRESSION

Used for predicting the likelihood of categorical values such as diagnosis or special outcomes. Can also be used for complex applications such as probability estimation; e.g., Probabilistic Graph Models (PGM).

COMMON USE CASES

- CDS and diagnosis tools; e.g., a model for diagnosing Chronic Obstructive Pulmonary Disease (COPD) might include historical and demographic information such as age, sex, smoking history, exposure to chemicals, and signs and symptoms such as coughing, dyspnea, and blood oxygen saturation, along with comorbidities such as bronchitis, diabetes, and lung cancer. Each of these factors would be represented as a node in the graph, with connections between nodes indicating a causal relationship. The presence or absence of a combination of these factors will influence the probability that a COPD diagnosis is applicable to the patient

IMAGE AND VIDEO UNDERSTANDING

WHAT'S THIS IMAGE ALL ABOUT? WHAT CAN YOU INTERPRET FROM THIS VIDEO?

Example: Cancer tissue analysis through image scans.

DEEP LEARNING

Deep learning systems using CNNs (Convolutional Neural Networks) have the capability to identify the hidden patterns inside images and videos. Deep learning algorithms have the capability to analyze images and what's in them, without explicitly telling them what is crucial in it. Deep learning models need to be trained by being shown millions of images, each labelled with the desired variable marked. This training process is computationally intensive but once its trained, it can be used to easily identify the patterns.

COMMON USE CASES

This is one of the hottest area for machine learning in medical science.

Some specific use cases include:

- Workflow monitoring and procedural compliance
- Patient safety monitoring
- Radiology image analysis — a lot of deep learning models are used to detect pneumonia, cancer and other chronic disease patterns inside scanned images. There are also efforts to integrate this capability into existing radiology detection systems like MRI, CT and ultrasound

SPEECH-TO-TEXT

WHAT IS THE CORRESPONDING TEXT FOR AN AUDIO STREAM?

Example: Conversion of physician dictation to text

HIDDEN MARKOV MODELS (HMM)

HMM is a statistical Markov model in which the system being modeled is assumed to be a Markov process with unobserved; i.e. hidden states. These models are best known for their application in reinforcement learning and temporal pattern recognition such as speech, handwriting, gesture recognition, part-of-speech tagging, musical score following, partial discharges and bioinformatics.

LONG SHORT-TERM MEMORY (LSTM)

LSTM models are similar to Convolutional Neural Networks described above. They can automatically learn what attributes of an audio stream are important for predicting what words they represent. Given sufficient data, which is readily available via online service providers such as Google and Baidu, it is possible to train LSTM models to accurately convert spoken language into text in almost any language. This technology has become the state-of-the-art for spoken language understanding applications and will likely play an increasing role in clinical transcription applications.

COMMON USE CASES

Classifiers are the most commonly used machine learning algorithms in all analytics applications, including healthcare.

Some specific use cases include:

- Dictation and clinical note prescription
- Voice controls for computer systems in the clinic and in the surgical theater
- Interpretation and automated documentation of clinical encounters,
- Call-center resources and agent coaching to help call-center operators provide appropriate information and resources to patients or members
- Patient coaching — applications are being developed in which patients are given

CLUSTERING

CAN THE DATA BE GROUPED INTO THE NATURAL CATEGORIES OR BUCKETS?

Example: Can I logically group data based on similarity?

UNSUPERVISED CLUSTERING

Unsupervised clustering algorithms such as K-means can logically group data based on their mutual distance and create centroids to create groups. This process is based on different parameters of data which can successfully divide the data into logical groups. This is a highly explainable algorithm which can be important in healthcare to explain how conclusions are reached.

HIERARCHICAL CLUSTERING

Hierarchical clustering involves creating clusters that have a predetermined ordering from top to bottom. For example, all files and folders on the hard disk are organized in a hierarchy. There are two types of hierarchical clustering, Divisive and Agglomerative.

COMMON USE CASES

- Grouping similar patients — In the care of complex or rare disease, it is valuable to understand how different treatments have worked on other patients in similar situations, but this information is only useful if the past patients are similar enough to the current patient to have predictive value. Clustering can be very useful to identify the most similar patients.

REFERENCES

1Obermeyer, Z. and Emanuel, E.J. (2016, September 29). “Predicting the Future—Big Data, Machine Learning, and Clinical Medicine.” New England Journal of Medicine, Vol. 375, p. 1216.

Demystifying Big Data and Machine Learning for Healthcare https://www.medcalc.org/manual/logistic_regression.php

https://www.kdnuggets.com/2016/07/support-vector-machines-simple-explanation.html

https://medium.com/@kangeugine/hidden-markov-model-7681c22f5b9

“Convolutional Neural Network.” https://en.m.wikipedia.org/wiki/Convolutional_neural_network Yarin, G. op. cit.

“Reinforcement Learning.” (n.d.) Wikipedia. https://en.m.wikipedia.org/wiki/Reinforcement_learning

A RICH OPPORTUNITY FOR DISCOVERY

In addition to the use cases outlined above, there are many other areas where machine learning is being applied for great benefit in healthcare including recommendation systems, Information retrieval, text mining and sophisticated information warehousing techniques. Healthcare is a rich domain for the technology owing to the dimension and propensity of information associated with it. The marriage of math, statistics and business application lends itself to many use cases for the betterment of patients and healthcare providers.

Apexon

APEXON IS A PURE-PLAY DIGITAL ENGINEERING SERVICES FIRM FOCUSED ON HELPING COMPANIES ACCELERATE THEIR DIGITAL INITIATIVES FROM STRATEGY AND PLANNING THROUGH EXECUTION.

We leverage deep technical expertise, Agile methodologies and data-driven intelligence to modernize systems of engagement and simplify human/tech interaction.

We deliver custom solutions that meet customers' technology needs wherever they are in their digital lifecycle. Backed by Goldman Sachs and Everstone Capital, Apexon works with both large enterprises and emerging innovators — putting digital to work to enable new products and business models, engage with customers in new ways, and create sustainable competitive differentiation.

info@apexon.com www.apexon.com

FEELING SOCIAL?