panalgo

How Machine Learning is Being Leveraged to Analyze Real-World Data

A Collection of Recent Research and Use Cases

Table of Contents

INTRODUCTION

Machine Learning: An Essential Addition to the RWE Toolbox

It's well known and often touted that real-world data (RWD) is growing at an unprecedented rate both in terms of volume and variety. As a result, the demand to crunch larger datasets and generate novel insights has intensified. With a competitive edge going to those who generate insights and answer crucial business questions faster, advanced analytic techniques, like machine learning (ML) are being employed more and more to maintain this advantage.

Decision-makers must create clear strategies for augmenting traditional analytics with ML-driven approaches—but the most commonly asked question is *how* to best use machine learning as a part of an overall analytics strategy. How will harnessing this advanced technology benefit companies in the long run? Will it become the new preferred analytics method? When is ML preferable to other analytics? And yet, there is no simple answer. Instead, it depends on the type and quantity of data available, and most importantly, the goals of the research. That's why it's important to understand how ML is being employed, and what its potential is—which is why we have put this library of ML use cases together.

Understanding Machine Learning

Machine learning is a type of artificial intelligence (AI) consisting of a collection of methodologies that focus on algorithmically learning efficient representations of data and then extracting insights from the data. It has consistently been gaining traction within the industry, especially within in the context of RWD.

There are two primary types of machine learning:

- Supervised ML algorithms label the data to tell the machine the specific patterns to look for.
- Unsupervised ML algorithms have no labels and seek out whatever patterns it can discover.

Positive Outlook for Machine Learning

In a recent survey of life sciences executives, 95% of respondents said they expect to use machine learning in the coming years to glean real-world evidence (RWE) from the growing volume of data. Additionally, respondents in the Panalgo 2021 Benchmarking Report reported a positive outlook on the benefits of machine learning and data science. Two-thirds (66%) of respondents indicated that outcomes research/health economics would be significantly improved with ML, while more than half of respondents reported significant improvements to trial protocol design/optimization (59%), market forecasting (57%), clinical trial recruitment (57%), disease identification/population sizing (56%), and evidence generation to support regulatory submissions (52%) with advanced analytics like ML.

Activities that Would be Significantly Improved if Machine Learning Could be Applied

To better understand the impact of machine learning on analyzing real-world data, we have compiled 12 recent use-case studies across healthcare, divided into seven major categories:

1. Identifying Meaningful Patient Subgroups
2. Prediction of Clinical Outcomes
3. Derivation of RWD-based Phenotyping Algorithm
4. Rare Disease Prediction
5. Optimizing Current Processes & Improving Workflows
6. Propensity Score Estimation
7. Prediction of Optimal Treatment

IDENTIFYING MEANINGFUL PATIENT SUBGROUPS

USE CASE #1

Identifying and evaluating clinical subtypes of Alzheimer's disease in care electronic health records using unsupervised machine learning

University College London, Health Data Research UK, Alan Turing Institute, & Amsterdam University Medical Centers

Published: *BMC Medical Informatics and Decision Making*

According to the Alzheimer's Association, on average, a person with Alzheimer's lives four to eight years after diagnosis, but can live as long as 20 years, depending on a number of factors. To better understand clinical outcomes in this patient population, researchers in this study used unsupervised machine learning clustering techniques and RWD to identify a meaningful subgroup of Alzheimer's disease (AD) patients associated with worse clinical outcomes. They found that a subset of female patients with younger disease onset and comorbid depression and anxiety are more likely to have worse clinical outcomes and recommend further investigation of the health outcomes for this patient subtype.

Approach

AD patients were identified in primary care electronic health records (EHR) from the Clinical Practice Research Datalink (CPRD) using a previously validated rule-based phenotyping algorithm. Comorbidities, symptoms, and demographic features were included, and four different clustering methods were evaluated (k-means, kernel k-means, affinity propagation and latent class analysis) to cluster AD patients. Clusters were compared on clinically-relevant outcomes and each method was evaluated using measures of cluster structure, stability, efficiency of outcome prediction and replicability in external data sets.

Results

There were 7,913 AD patients identified with a mean age of 82, and 66.2% were female. A total of 21 features were included in the analysis. Cluster patterns were identified to evaluate the different subtypes of AD. A consistent cluster was found in three of the four methods, which included mostly young females (43% between ages 42–73) diagnosed with depression and anxiety. This cluster had a quicker rate of progression compared to the average across other clusters.

Conclusion

While there were some inconsistencies in the different clusters and K-means performed, on cluster remained relatively consistent across three of the four methods. This particular subtype of AD could potentially be further explored, such as by determining whether the clusters contain genetic or neurophysiological differences.

PREDICTION OF CLINICAL OUTCOMES

USE CASE #2

Predicting hypoglycemia in critically ill patients using machine learning and electronic health records

Harvard University & Beth Israel Deaconess Medical Center & MIT et al.

Published: Journal of Clinical Monitoring and Computing

There is great optimism that the application of artificial intelligence (AI) can provide substantial improvements in all areas of healthcare from diagnostics to treatment. By evaluating 15 machine learning models, researchers developed a predictive model to predict hypoglycemia. The resulting machine learning model had strong discrimination as well as calibration and performed better than the baseline logistic regression model.

Approach

Using the eICU Collaborative Research v2.0, researchers extracted data and analyzed 69,736 patients who were admitted to the hospital and had fewer than two blood glucose readings. Fifteen machine-learning models were evaluated in their performance.

Results

The best-performing model was the XGBoost model which ran more efficiently and had high sensitivity, meaning its algorithm can identify most of the patients who were experiencing hypoglycemia.

Conclusion

As machine learning tools continue to advance and build predictive power, some may be able to be used in hospital settings to alert clinicians of patients at high-risk of hypoglycemia in real-time.

USE CASE #3

Early prediction of hemodynamic interventions in the intensive care unit using machine learning

Massachusetts Institute of Technology & Phillips Research North America

Published: *Critical Care*

Hemodynamic instability occurs when there's abnormal or unstable blood pressure, which can cause inadequate blood flow to the organs. Symptoms may include chest pain, abnormal heart rate, or low blood pressure.

In a recently published study in BMC, researchers developed a machine-learning based index for identifying hemodynamic instability in critically ill patients, rather than relying on basic vital signs such as heart rate, blood pressure, and shock index. The index may help clinicians catch instability and deliver intervention earlier.

Approach

A model was developed using the eICU Research Institute (eRI) database, which is based on ICU admissions of adult patients from 2012-2016. A total of 208,375 patients met criteria for the study, with 32,896 patients who experienced at least one instability event.

Results

The Hemodynamic Stability Index (HSI), a model developed to identify hemodynamic instability in critically ill patients, performed significantly better in comparison to relying on basic vital signs.

Conclusion

This model is a real-time measure for predicting hemodynamic instability. It can make predictions for future interventions with confidence. It will be critical in real-world situations and will provide actionable prompts based on impact score.

USE CASE #4

Predicting response to tocilizumab monotherapy in rheumatoid arthritis: A real-world data analysis using machine learning

Chalmers University of Technology
Brigham and Women's Hospital,
Genentech, MIT et al.

Published: *The Journal of Rheumatology*

Rheumatoid arthritis (RA) is a chronic autoimmune disease that affects more than 1.3 million people in the United States. In this study, researchers compared two methods of prediction: utilizing machine learning and high dimensional RWD vs. obtaining a risk score using randomized controlled trials (RCT) data to predict response to tocilizumab monotherapy in rheumatoid arthritis.

Approach

Patients in the Corrona RA registry who met the study criteria were observed to determine remission status at 24 weeks. Then, researchers compared the performance of remission models based on variables previously determined in their work in RCTs.

Results

The patients who reached remission on TCZm by 24 weeks were 12% in RWD and 15% in RCTs. Discrimination measured by AUCROC improved from 0.69 to 0.72 when the risk score developed using RWD.

Conclusion

The prediction scores derived in RCTs discriminated patients in RWD almost just as well as with RTCs. By retraining the models on RWD, the prediction scores improved even further.

USE CASE #5

Machine learning study of real-world data: Predicting inpatient relapse in multiple sclerosis patients using first-line disease modifying therapies

Panalgo Original Research:
Zeynep Icten, PhD, Mark Friedman, MD, Joseph Menzin, PhD.

Published: *2021 ISPOR presentation*

Life sciences leaders believe machine learning analytics would significantly improve outcomes research, market forecasting, disease identification/population sizing, and evidence generation to support regulatory submissions. When looking specifically at multiple sclerosis (MS), understanding the drivers of relapse can inform improved management of the disease. In this case study, we use IHD Data Science to identify predictors of inpatient MS relapse.

Approach

The IHD Data Science module provided the ability to rapidly and efficiently test different models enabling researchers to measure performance across six different model types: XGBoost, Random Forest, Neural Network, Regularized Logistic Regression, Support Vector Machine, and Logistic Regression. Features included demographics, comorbidities, concomitant medications, healthcare resource utilization, disease modifying therapies (DMTs), route of administration and proportion of days covered for DMTs.

Results

The XGBoost ML model had the strongest ability to predict inpatient MS relapse compared to all models tested. The IHD Data Science module allowed researchers to assess the patient's full diagnosis history and evaluate a myriad of factors that may be related to MS relapse, including previous inpatient or emergency room visit with an MS diagnosis, the number of MS related encounters, and more. The model demonstrated that MS patients are more likely to have a relapse if they 1) have 30 or more unique comorbidities, or 2) have a previous emergency room visit with an MS diagnosis and 10 or more previous MS related encounters or 3) have 20 or more previous MS related encounters.

Conclusion

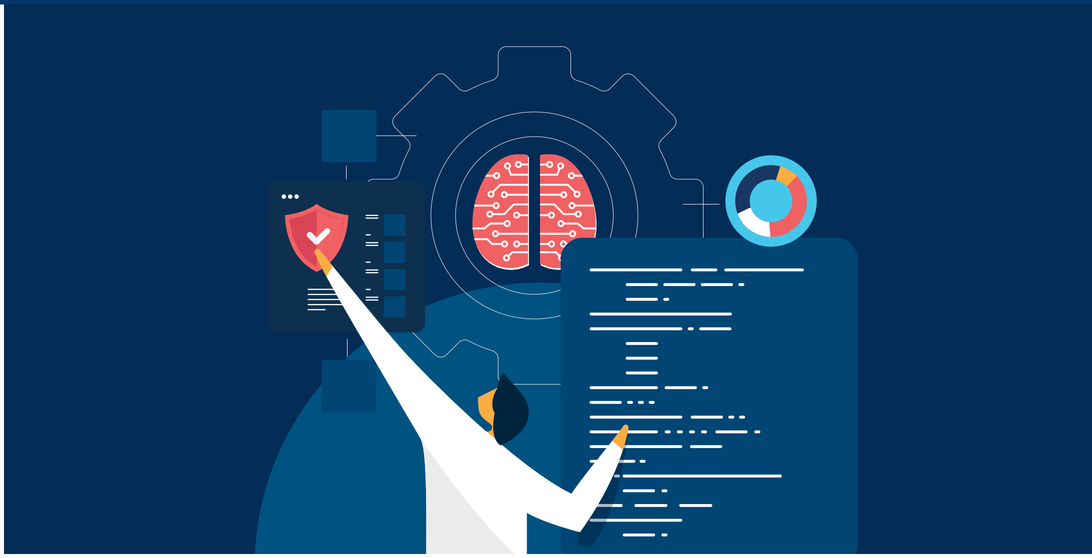
Understanding the drivers of MS relapse beyond well-known factors allows life sciences to more accurately target patients for therapy and gives providers the ability to identify novel subgroups of patients who are at high risk for relapse and improve disease management.

USE CASE #6

Electronic phenotyping of health outcomes of interest using a linked claims-electronic health record database: Findings from a machine learning pilot project

Government Health and Human Services, IBM Watson Health, Bethesda, Maryland, Food and Drug Administration, Silver Spring, Maryland, Harvard Medical School and Harvard Pilgrim Health Care Institute, et al.

Published: *Journal of American Medical Informatics Association*



The Food and Drug Administration (FDA) Sentinel Active Risk Identification and Analysis System use claims-based algorithms to find occurrences of health outcomes of interest (HOIs) for medical product safety assessment. This study used machine learning to demonstrate the potential ability of developing a claims-based algorithm to predict an HOI in electronic health records (EHRs).

Approach

Researchers used the 2015-2019 IBM MarketScan Explorys Claims-EMR Data Set and linked administrative claims and EHR data at the patient level. The HOI rhabdomyolysis, defined by EHR laboratory test results, was predicted by using claims-based predictors and machine learning techniques. These included logistic regression, least absolute shrinkage and selection operator (LASSO), random forests, support vector machines, artificial neural nets, and an ensemble method (Super Learner).

Results

Model performance varied considerably across techniques. The area under the receiver-operating characteristic curve exceeded 0.80 in most model variations. The Super Learner ensemble model performed considerably better than a human model without adjustment for class imbalance.

Conclusion

It is feasible to use machine learning to predict an EHR-derived HOI with claims-based predictors, and the models can be adjusted for various uses such as surveillance, identification of cases for chart review, and outcomes research.

USE CASE #7

Effectiveness and safety of rivaroxaban versus warfarin among nonvalvular atrial fibrillation in patients with obesity and diabetes

University of Maryland School of Medicine & Janssen Scientific Affairs

Published: *Journal of Diabetes and its Complications*

According to the Centers for Disease Control and Prevention (CDC), more than 454,000 hospitalizations with atrial fibrillation (AF) as the primary diagnosis happen each year in the United States and contributes to about 158,000 deaths each year. Researchers used a machine learning algorithm to impute obesity in a claims database study in order to evaluate the effectiveness and safety of rivaroxaban compared with warfarin in nonvalvular atrial fibrillation patients. They concluded that the use of the algorithm substantially increased the sample size and study generalizability.

Approach

Patients 18 years and older were identified from a healthcare claims database with the following criteria: newly initiating rivaroxaban or warfarin, ≥1 medical claim with a diagnosis of atrial fibrillation (AF) obesity determined by validated machine learning algorithm, and ≥1 claim with a diagnosis of diabetes or for antidiabetic medication. Additionally, there was a comparison for risk of stroke/systemic embolism (SE) and major bleeding across treatment cohorts.

Results

The risk of stroke/SE was much lower in the rivaroxaban cohort compared to the warfarin cohort. Additionally, the risk of ischemic and hemorrhagic strokes was significantly reduced with rivaroxaban compared to warfarin, but not SE. Finally, the major bleeding risk was similar across the treatment cohorts.

Conclusion

In patients who had NVAF in addition to comorbidities of obesity and diabetes, rivaroxaban was associated with lower risks of stroke/SE and similar risk of major bleeding when compared to warfarin.

USE CASE #8

Revisiting performance metrics for prediction with rare outcomes

NYU & Harvard Medical School & MGH & Stanford University

Published: *Statistical Methods in Medical Research*

Machine learning uses various statistical techniques and advanced algorithms to predict the results of healthcare data more precisely. This study provides an evaluation of using machine learning to predict rare outcomes, highlighting key areas to consider including the dangers of relying on a single measure for prediction.

Approach

Researchers used multiple algorithms and evaluation metrics—which had previously performed well with rare outcomes—to predict post-surgery mortality among patients in the Massachusetts Data Analysis center who received at least one aortic valve replacement.

Results

It was discovered that high accuracy could be met with a very low true positive rate (TPR) for predicting both short-term and long-term mortality following AVR.

There was a high accuracy for all algorithms, with false positive rates at <1%, but true positive rates were <7%. In the simulations, the results were relatively similar, with a high AUROC associated alongside the low true positive rates.

Conclusion

Researchers recommend that future medical studies focus more on prediction rather than accuracy as the essential metric to have a better understanding of the algorithm. Thus, the need for more developments of algorithms for targeting rare outcomes is needed and has a promising future when considering all variables.

USE CASE #9

Using machine learning to examine drivers of inappropriate outpatient antibiotic prescribing in acute respiratory illnesses

The Centers for Disease Control and Prevention & IQVIA

Published: *Infection Control & Hospital Epidemiology*

In 2016, nearly one in four antibiotic prescriptions were unnecessary. In a recent study published in Infection Control and Hospital Epidemiology, researchers from the CDC and IQVIA demonstrated the ability of machine learning to identify the drivers of inappropriate antibiotic prescribing. Evaluating these drivers helped the CDC target policy intervention.

Approach

Researchers identified visits and inappropriately prescribed antibiotics using IQVIA's Medical Claims Data set and Longitudinal Prescription Data. Using a machine-learning model, they evaluated drives contributing to the inappropriate prescriptions such as age, sex, training, state, setting, and gender mix.

Results

The machine learning tool helped to detect variations in location and other drivers that were strong predictors of inappropriate antibiotic prescribing. Complex relationships across these drivers were found, particularly in Alabama. It was noted that the highest average of inappropriate prescribing happened among NPs and PAs among pediatricians. Finally, the urgent care setting was the strongest driver.

Conclusion

Using machine learning tools helped to evaluate a wide range of features contributing to the inappropriate antibiotic prescribing. These tools might be expanded to be used in other specialties and settings within healthcare.

USE CASE #10

Improving patient flow during infectious disease outbreaks using machine learning for real-time prediction of patient readiness for discharge

University of Oxford &
Oxford University Hospitals
NHS Foundation Trust

Published: *PLOS ONE*

Leading up to the COVID-19 outbreak, approximately 728,000 medical and surgical hospital beds were available to the public, or 2.2 beds per 1,000 people. However, only 36 percent of these beds were unoccupied on a typical day, leaving around 0.8 unoccupied beds per 1,000 people. In response to a shortage of hospital beds during the COVID-19 pandemic, Oxford researchers developed and validated a machine learning-based modeling technique to identify the patients with the highest likelihood of discharge readiness. The research published in *PLOS ONE* identified a model that may provide support and insights to clinicians for making hospital discharge decisions during periods of public health crises when resources are limited.

Approach

Data from the Oxford University Hospitals' Electronic Health Record was used to identify patients' real-time readiness to be discharged within a span of 24 hours. Predictions were made for patients in order of likelihood of discharge, and those with the highest ranking were candidates and were expected to be screened to be correctly deemed ready for discharge.

Results

Predictions using machine learning performed best in predicting a patient's discharge of planned rather than emergency admissions. Models generally did not indicate early discharge for patients, boosting confidence that they would not make an unsafe prediction.

Conclusion

Using this model will continue to improve the efficiency of overall patient flow in and out of hospitals. It would also help to predict real-time discharge for individual patients.

PROPENSITY SCORE ESTIMATION

USE CASE #11

Global pharmaceutical company uses Panalgo's data science module to leverage real-world data to design and execute prospective observational study

Published: *Panalgo Case Study*

As life science companies continue to expand their capabilities within RWD, they rely on data science modules to help study design and execution. For a new medication, a global pharmaceutical company was organizing a study that would help inform the design and execution of the product's prospective observational study–comparing patients exposed to a new product to those following the standard of care. Specifically, the company needed to better understand the patient population that would best fit the study and minimize uncertainty in its execution through studying real-world data.

Approach

To understand the patient population needed to enroll in an observational study and minimize uncertainty, appropriate patient sample size, the case-to-control ratio, and the propensity score match rate were determined to ensure the patient sample population was reflective of the real world.

Results

Panalgo accelerated project from months to approximately six weeks and set stage for increased productivity for future analytics projects. Panalgo helped the company win an internal innovation award for its cutting-edge use of analytics for observational study design.

Conclusion

IHD Data Science module enabled the team to test multiple predictive models, including novel machine learning models, to identify and choose the best model for the study's needs and validate that chosen model. Panalgo solutions team expertise expedited and optimized the project.

PREDICTION OF OPTIMAL TREATMENT

USE CASE #12

Predicting optimal treatment regimens for patients with HR+/HER2- breast cancer using machine learning based on electronic health records

Eli Lilly & Company, Indianapolis & Vienna

Published: *Journal of Comparative Effectiveness Research*

Breast cancer is the most frequently diagnosed cancer in women worldwide. In 2020, 276,480 new breast cancer cases are estimated to be diagnosed and a staggering 42,170 deaths are estimated to occur in the US alone. This study aimed to predict optimal treatments maximizing overall survival (OS) and time to treatment discontinuation (TTD) for patients with metastatic breast cancer (MBC) using machine learning methods on electronic health records.

Approach

The patient populations for this study consisted of adult females with HR+/HER2- MBC on first- or second-line systemic therapy. Random survival forest (RSF) models were used to predict optimal regimen classes for individual patients and each line of therapy based on baseline characteristics.

Results

The RSF models predicted optimal treatments prolonged OS and TTD when compared against nonoptimal treatments in patients with HR+/HER2- MBC on either first- or second-line therapy. The models also predicted optimal treatments when compared against other treatments in the real world. Key variables such as age, weight, date of metastatic diagnosis, ER status, PR status, Breasts CAncer gene (BRCA) status, practice region of the US, practice type, race, cancer stage at initial diagnosis, ECOG PS, first-line regimen class, and TTD of first-line therapy were crucial to help with these predictions.

Conclusion

RSF may help inform optimal treatment choices, assist in physician decision-making, and improve outcomes for patients with HR+/HER2- MBC.

Summary

As the data grows in complexity and volume, we need to shift towards the tools and approaches that work best with such high dimensional data to generate the best insights within RWD. Machine learning has been gaining traction year over year because of its ability to connect relationships within the data, especially when complex and nonlinear. Its ability to make predictions and generate novel insights is why it is a valuable addition to the RWE toolbox.

About Panalgo

Panalgo provides software that streamlines healthcare data analytics by removing complex programming from the equation. Our Instant Health Data (IHD) software empowers teams to generate and share trustworthy results faster, enabling more impactful decisions. Our Data Science Module combines the speed of IHD with the power of machine learning.

CLICK HERE TO LEARN MORE