2018

Systemic Lupus Erythematosus Symptom Severity Prediction Using a Recursive Neural Network

Katherine G. Skocelas
Grand Valley State University

Systemic Lupus Erythematosus Symptom Prediction Using Recursive Neural Network

By
Katherine G. Skocelas
December, 2018

Systemic Lupus Erythematosus Symptom Prediction Using Recursive Neural Network

By
Katherine G. Skocelas

A project submitted in partial fulfillment of the requirements for the degree of
Master of Science in
Computer Information Systems

at
Grand Valley State University

December, 2018

Table of Contents

Abstract

Systemic Lupus Erythematosus (SLE) is a chronic autoimmune disease that causes the immune system to attack the body's own connective tissues and organs. Humans have difficulty predicting SLE symptom severity levels because of the complex interactions of disease trigger exposure levels over time. To address this issue, we constructed a novel machine learning solution that generates a model capable of predicting SLE symptom severity levels with 8.3-19.9% average error. It does so by inputting trigger exposure levels into a recursive neural network and training them with a unique method that continually turns training on and off based on the maximum error each day. This allows the RNN to learn SLE flare activity without overtraining remission activity, thus maintaining a greater degree of plasticity. Models trained in this fashion performed 3.5-5% better on average than those trained via the standard method. Future areas of work include replicating these results with a large patient training data set and evolving the model to predict disease trajectory.

Introduction

Systemic Lupus Erythematosus (SLE) is a chronic autoimmune disease that causes the immune system to attack the body's own connective tissues and organs. It is a serious and potentially life-threatening illness that affects roughly five million people around the world (Lupus Foundation of America, 2018). Humans have difficulty predicting SLE symptom severity levels because of the complex interactions of disease trigger exposure levels over time. SLE symptom severity levels continuously fluctuate in response to the individuals' sensitivity and exposure to disease triggers, such as activity and stress level, UV exposure, amount of sleep, illness and injury. To manage their symptoms, SLE patients must understand and limit their exposure to these triggers. Because of their complex interactions over time, however, this is not a simple task.

For example, an individual may be able to safely tolerate three hours of UV exposure when well-rested and relaxed. If they have other compounding trigger exposures, however, the same about of UV exposure may cause them to develop a rash. Additionally, the same amount of trigger exposure can cause different reactions depending on the amount of exposure experienced the previous days. This is not only true for the same trigger, but others as well. This means that the same three hours of UV exposure when relaxed could cause a reaction if the previous day the individual was extremely stressed.

This complexity can make SLE feel unpredictable, unmanageable and frightening to sufferers. "With symptoms that come and go, disease flares and remissions, and the uncertainty of what each day will bring, it's normal to experience feelings of unhappiness, frustration, anger, or sadness" (Lupus Foundation of America, 2018). This emotional tole can become debilitating. Studies show that between 15 and 60 percent of people with a chronic illness like SLE will experience clinical depression during the course of their disease (Lupus Foundation of America, 2018).

A tool that is able to predict SLE symptom severity levels with a greater accuracy than humans is one way to reduce this burden. By identifying and illuminating patterns in trigger exposure and symptom severity levels, such a tool would also improve SLE patients' ability to manage their symptoms. This would have profound and wide-ranging effects. Better symptom management has been shown to increase patient quality of life, reduce healthcare costs, and improve disease prognosis (Lupus Foundation of America, 2018).

Background and Related Work

Machine learning is an emerging field in computer science that seeks to solve complex problems like the one posed by SLE. Because there are many trigger exposure levels that interact to result in many symptom severity levels, the problem is well suited to a specific type of machine learning architecture known as an artificial neural network (Fausett, 1994).

Artificial neural networks are a type of computing system inspired by the biological neural networks that constitute animal brains (Fausett, 1994). They are comprised of layers of nodes with weighted connections between them that are meant to immitate neurons (Fausett, 1994). In their simplest

execution, data is loaded into the input layer nodes, filtered through the network, and outputed via the output layer nodes (Fausett, 1994).

There are many unique forms that artificial neural networks can take to improve their performance in different problem domains. In one structure, called a recursive neural network (RNN), the previous step's output values are reentered into the net as input values for the next step (Fausett, 1994). This is useful for problems like SLE symptom prediction in which previous outputs influence future predictions.

Context units are another type of structure that allow neural networks to account for history-sensitive variance (Fausett, 1994). A special type of input node, these units receive their data from the previous steps' hidden layer nodes and reinsert it into the net during the next step. This injects historical input data into the equation (Fausett, 1994).

We hypothesized that an artificial neural network could successfully predict SLE symptom levels because it has been used successfully for other disease symptom prediction problems (Wu, Roy, & Stewart, 2010). The specific form of an RNN with context units was selected due to its proven ability to account for historical input and output data in a variety of problem domains (Fausett, 1994). This allows it to account for the influence of historical trigger exposure and symptom severity levels in its predictions.

Existing medical machine learning studies have used a similar time series approach, such as Wiens, Guttag, and Horvitz's (2012) work predicting patient risk of hospital-acquired infection. The current body of work is limited, however, in its application to SLE.

Existing machine learning studies addressing SLE treatment are focused on clustering problems, such as identifying the genes responsible for SLE (Armañanzas et al., 2009) and risk factors for specific adverse SLE outcomes (Ravenell et al., 2012). These studies utilize quantitative data, such as DNA, that is collected and analyzed in a laboratory setting. We did not find any SLE studies that utilized or predicted subjective patient data such as fatigue and pain levels.

This deficit is likely due to the difficulties associated with the accurate elicitation and normalization of subjective health data for use in computer algorithms. While gathering daily symptom level reports, we have found that factors such as emotional state and personality greatly impact reported values. This negatively affects the predictive ability of models that use the data. The inability to correct for this issue is a significant weakness in the current body of knowledge, because it greatly limits what aspects of SLE can be effectively studied.

Program Requirements

To begin to fill the existing knowledge gaps, we have designed, implemented, trained and tested an RNN for SLE symptom prediction based on historic disease trigger exposure and symptom severity levels. We hypothesized that this was possible because machine learning has been used successfully for other disease symptom prediction problems (Wu, Roy, & Stewart, 2010). The model produced shows that, though humans struggle to predict SLE symptoms, they are based on a pattern of trigger exposure interactions that is machine learnable.

The model is able to memorize a year of artificial SLE data to a roughly 20% degree of accuracy. This proves that, though humans struggle to predict SLE symptoms, they are based on a pattern of trigger exposure interactions that is machine learnable. Interestingly, the model memorizes actual patient data sets more accurately than it does the artificial data sets. This suggests that there may be conflicting data in the artificial sets that was accidentally injected in the construction process. If this hypothesis is true, it means that the problem is even more well suited to a machine learning solution than is evident by our results.

The model produced by the standard train/validate process, however, is too rigid to perform well on diverse and fluctuating SLE presentations. SLE includes hallmark disease flare and remission cycles, and the remission periods were being over-trained though the flares were under-trained. This led us to construct a second, novel training method for the RNN that is more performant in this disease domain. It continually turns training on and off based on the maximum error each day. When triggered on, the RNN retrains over the past 30 days before continuing training new incoming data. In this way, the RNN is able to learn the SLE flare activity without overtraining the remission activity.

Because SLE is a rapidly shifting disease, the new model's rate of change is also significantly increased, allowing it to react quickly to a slowly incoming data set. In our study, models trained in this fashion performed 3.5 to 5% better on average than those trained via standard memorization. Although a synthesized training data set was used because of data access limitations, the models generated were accurate on both small real patient data sets. In the future, we would like to replicate these results with a large SLE training data set.

Implementation

The RNN built is composed of a five-node input layer with two recursive inputs and ten context units, a ten-node hidden layer, and a two-node output layer (see Appendix A). The patient's daily stress level, activity level, UV exposure, hours of sleep, and illness or injury level are entered into the network. The model outputs predicted fatigue and rash levels. The actual values of these symptoms are then entered and used the following day as the recursive inputs. If it is in training mode, error is calculated and backpropagated through the RNN.

Training a machine learning algorithm requires large data sets (Fausett, 1994). Because of limited access to patient data, we built large training and testing data sets using artificially constructed patient data. The accuracy of these sets was reviewed by an SLE patient and verified in the model via their predictive capacity on two small SLE patient data sets.

The original concept for this project was an iOS application for SLE patients that would allow them to enter trigger exposure levels and get predicted symptom levels. Within the first weeks of the project, however, we realized that this prediction could only be accomplished via machine learning due to the complex interaction of trigger exposure levels over time. Never having studied machine learning before, we decided to implement the RNN without external machine learning libraries so that we could better learn and understand what it was doing and to accommodate our novel training model.

External machine learning libraries such as TenserFlow were considered and may be used in subsequent prototypes. Independently building the model, however, had significant academic value. It allowed us to apply the new information we were learning in a hands-on way, and it revealed points of confusion. For example, several misunderstandings were made evident when we attempted to implement the backpropagation method, including how the values were being calculated and where the signals were being sent. These would not have come to light and been corrected if we had simply made a function call to an external library.

Results, Evaluation, and Reflection

This project is an initial prototype into a complex and previously unexplored problem space. It faced substantial time and data availability limitations, and it was a stretch of our academic abilities. As such, the results are significant for several reasons.

First, the model's ability to memorize the artificial and SLE data sets indicates that SLE symptom severity levels are based on a pattern of trigger exposure interactions that is machine learnable. This opens the door to further machine learning research into their relationship and prediction. This discovery was able to be made despite limited access to SLE patient data because we spent considerable time building and testing artificial data sets that could be used to train and test the RNN.

Second, we were able to construct a novel training method that is 3.5-5% more effective than the standard method because we took the time to construct the RNN from scratch instead of using existing libraries. This allowed us to learn exactly how the RNN was working (despite being new to machine learning research), which revealed the issues with the standard model. Further work is needed to reduce the average error as much as possible.

Finally, the end result represents a profound amount of self-directed independent learning. Undertaking it was our first exposure to machine learning concepts, and yet we were able to construct a working prototype that is reasonably accurate within the space of a single semester.

If enough patient data had been available to train and test the RNN without artificial data sets, and the same results were achieved, this would be publication-worthy work. It pushes the bounds of SLE machine learning research and proves that further exploration is likely to be fruitful. Most importantly, it provides a working predictive model which SLE patients can use to manage their symptoms.

Conclusions and Future Work

The most valuable outcome of the model is the actionable information it provides to physicians and patients. Though still in its infancy, it already shows promise for symptom management. The next step is replicating the experiment only using real patient data. Large volumes of data from a wide variety of SLE

patients is necessary to improve the model's accuracy and plasticity. We suggest using data from at least 50 different SLE patients. Specific attention should be given to increasing the model's ability to recognize disease flares. It is important to also test the full range of disease triggers and symptoms in the RNN to form a complete real-world model. Such data is available through existing medical and pharmaceutical organizations currently engaged in SLE research.

In addition to improving the model, we are working to evolve it to not only predict outcomes but also direct patients using behavior modification feedback, guiding them into a desired disease state. Fox, Ang, Jaiswal, Pop-Busui, and Wiens' (2018) recently proved in a similar disease symptom trajectory problem that machine learning is effective for multi-step forecasting, indicating that this is likely possible for SLE as well. Using this disease path prediction, future programs could direct patients to avoid specific trigger exposures when they are straying from the desired disease course.

If successful, this research would be life changing for patients with SLE and similar disorders. Behavioral coaching for chronic disease self-care is financially out of reach for most patients. An accessible digital version, such as a machine learning-powered app, would allow for greater access to this type of care and improved outcomes.

Such an application would not be useful, however, if it did not come with a pretrained model. Further research should consider the ways in which a model or set of models can be constructed for initial use, then honed to the specific disease presentation of each user. This would likely involve machine learning research examining the possible identification of unique SLE subtypes.

Future work should also consider ways in which the need for manual data entry can be reduced without reducing the model's predictive capacity. This may include the automatic uploading of medical test results, activity monitors, and related health information. Ranking the usefulness of input variables and eliminating those with the least influence may also help reduce the burden on users.

Finally, this project revealed significant limitations caused by the use of subjective health data, such as fatigue and pain levels, in machine learning. Further work should examine the accurate elicitation and normalization of subjective health data for use in computer algorithms. Detecting and accounting for these issues could improve the predictive abilities of models that rely on the data, like the one described here.

Bibliography

Armañanzas, R., Calvo, B., Iñaki, I., López-Hoyos, M., Martínez-Taboada, V., Ucar, E., …& Zubiaga, A. (2009, May 9). Microarray analysis of autoimmune diseases by machine learning procedures. *IEEE Transactions on Information Technology in Biomedicine, 13*(3), 341-350.

Fausett, L. (1994). *Fundamentals of Neural Networks: Architectures, Algorithms, and Applications.* Upper Saddle River, NJ: Prentice Hall.

Fox, I., Ang, L., Jaiswal, M., Pp-Busui, R., & Wiens, J. (2018). Deep multi-output forecasting: Learning to accurately predict blood glucose trajectories. *24th ACM SIGKDD Conference on Knowledge Discovery and Data Mining.*

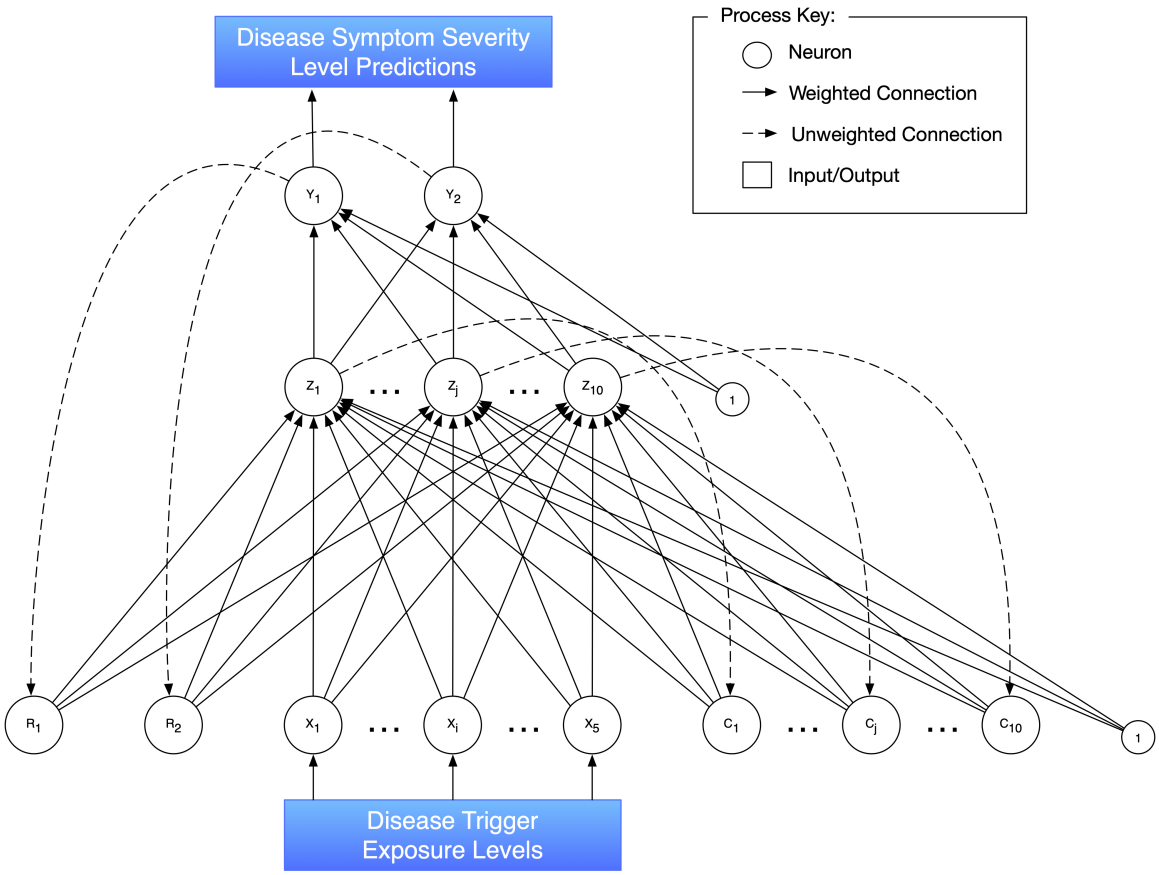
Lupus Foundation of America. (2013, July 31). *National Resource Center on Lupus*. Retrieved from Lupus.org: https://www.lupus.org/resources

Ravenell, R., Kamen, D., Fleury, T. J., Spence, D., Hollis, B. W., Janech, M. G., …& Almeida, J. S. (2012). Premature Atherosclerosis Is Associated With Hypovitaminosis D and Angiotensin-Converting Enzyme Inhibitor Non-use in Lupus Patients. *The American Journal of the Medical Sciences*, 268-273.

Wiens, J., Guttag, J., & Horvitz, E. (2012). Patient risk stratification for hospital-associated c. diff as a time-series classification task. *Proc. Advances in Neural Information Processing Systems.*

Wu, J., Roy, J., & Stewart, W. (2010). Prediction modeling using EHR data: Challenges, strategies, and a comparison of machine learning approaches. *Medical Care, 48*(6), S106-S113.

Appendix A



Architecture of 5 input, 2 output neural network with input units (X1-X10), recursive input units (R1-R2), context units (C1-C10), hidden units (Z1-Z10), & output units (Y1-Y2).