FINAL Assessment

Section 1: Agent-Based Models

Introduction to Agent-Based Models

The model one speaks about is called agent based model which explains how agents in a given situation make decisions and interact with one another. These are the best models for use where the micro-behaviors of the actors within the context result in emergent macro behaviors. ABMS have been used in many fields such as epidemiology, sociology or economics to model processes like disease spread, social relationships or market behavior. In this section, we focus on two add enhancements to two types of agent-based model: the discrete evolution model and the model of social dynamics. The first model contrasts aiming for a better estimation of parameters from one real data set whereas the second model furthers its design to make them usable and correct.

Part 1: Discrete Evolution Model

The discrete evolution model is intended to explain how the variant frequencies change and how they do so over time. Earlier on, the model we discussed did not make use of any inference function which could allow parameters to be predicted from real data samples. So, the task was to modify the model to enable the inclusion of such a function where the specific required assumptions are sampled from the data in section1_data1.csv which looks at the time steps and frequencies of variant 1 of a binary variant process at point A.

Enhancements and Methodology

The supplementing step function which extends the discrete evolution model has been implemented through the spline fitting function. In this case, UnivariateSpline suffices for the types of activities that require the simulation of the time distributin of the frequencies of variants. The spline is chosen in order to detect gradual trends in the movement of variant frequencies because it is important for knowing how these frequencies shift slowly over time. Thus a value of s = 0.5 is selected that is expected to retain the important features of the data while achieving a reasonably smooth curve so as to avoid overly fitting it to the data.

Then spline model is applied to obtain the posterior estimate of the variant frequencies mutations for the specific moments in time. This means the mean frequency of variant 1 and using 80% HDI as the lower interval of probability. From the range specified by the HDI, it tells us what part would probably characterize the true frequency and thus suggests this as being responsible for some of the sampling error. This allows not only predicting the occurrence of variant 1 but also predicting the confidence which can be anticipated in those predictions using the credible interval.

Results and Discussion

The plot presented in the above figure indicates the mean of the variant 1 over time and an 80% HDI is shaded around it indicating the results from using the enhanced model. The graph shows the spread of the variants accompanied by their decreasing numbers with both the spread and the decrease increasing, reaching a peak, and going back down, which indicates the existence of a constantly developing process of variants. Similar patterns such as these have been documented in other real-world instances, namely, where a variant, or an ‘innovation’, goes through a period of growth before reaching a plateau or decline.

Equally significant are the posterior estimates of the individual parameters. The mean frequency of variant 1 at every time point qualifies as one frequency that is therefore the average of values that were found while HDI is the extent of the credible interval and enables one to situate the values found in the place of variability. Now, there are certain periods where the model goes as far as being deemed accurate hence one can easily point out what those are: it is simply because the confidence levels were high for such intervals. The estimates regarding the individual parameters have provided the additional information which is about the probability of their true value lying in a certain specified range. In this instance, it is the use of variant 1, the mean amount at every time interval was used as the average frequency and the HDI was the average. Areas of confidence or belief are demarcated by these intervals, wherein the confidence is that the estimates are not widely off as they are narrow.

The spline model captures temporal changes in variant frequency yielding satisfactory results even when the processes in this domain become involved” Therefore, these changes to the model not only strengthen parameter inference with real data, but at the same time this feature bestows excellent credibility on the mathematical framework to analyze the behavior of the variants. Therefore, the Patel and Yu model attempts to regress the spread of a variant through the introduction of an inference function, and can as well be applied largely in the general study of variants.

Part 2: Social Dynamics Model

Discreet Evolution Model of Social Systems The last approach is the social dynamics model which is the generalization of the discrete evolution model that incorporates two versions of the model indicating where incidence occurs essentially by age. The earlier model contained the picture of only one variant which for sure explains the reason why it didn’t incorporate all of the social environmental factors. The task was to incorporate both variants in the model and define the initial levels of these two variants as a parameter unknown but which can be generated through inference.

Methodology

We used binary vectors in an agent to code for possession of fewer variants between the two variants, so that for instance (0, 1) denote that the agent possesses variant 2 but not variant 1. In this way, more detailed models of agents and their behavior in action and even behavior in interaction can be constructed, including a cooperative as well as an antagonistic action amongst variants.

The value of initial prevalence (init_p) of both variants has been made to an inferable parameter so that the model provides a better approximation of the realistic situation. Init_p for the reason, was previously constant at 0.1 but now it is a variable parameter which is to be deduced in the course of inference. This alteration allowed greater freedom with respect to the initial conditions and allow the model to be properly adjusted to reflect actual circumstances that may be encountered in practice.

Variants are spread over 50-time steps and the agent based model tries to estimate the optimal value of some criterion function that measures the distances between the simulated and target distributions. This process of optimization results into an initial prevalence estimate inferred of 0.10 and so there is remarkable agreement between model predictions and the data.

Results and Discussion

The results of the altered social dynamics model are presented in the form of a bar graph below depicting both target and simulated purposes of PSI by age variable. Regarding calibration, it is quite high, especially, for variant 1, The simulation means shifted to the location not far from around the target distributions. The forecasts of measured base-line prevalence are in line with a model fitting reality, and so it helps to treat init_p as a variable parameter.

The innovations introduced to the social dynamics model add an element of elasticity that increases the possibility for data set types. The model is able to have other settings when init_p is set to be inferred which also increases the model's forecast capability. This property is crucial since the research topic is concerned with more complex systems, since it is well known that no simple static systems or indeed their initial states and the processes acting on them elements are likely to be fixed over time.

The fixed approximate variation that emerged between the dropped variables and the original model explains the most significant conclusion that was drawn from the study. These modifications lead to increased precision and we notice the difference for the first variant, I believe that the alterations themselves are the reasons behind these discrepancies. In this manner, the analysis of variants' relations to each other explains that the new model discloses linkages and rivalries that standard Social Identity Theory does not take into account.

Conclusion

All two sections of the task, the agent-based models usage in this case allows one to come up with effective ways of doing dynamic simulation. The discrete evolution model is particularly effective in reconstructing the frequencies of variants over time and for the social dynamics model, prevalence across cohorts. Such models provide useful insight into how variants are changing and may have applications in the study of epidemiology and social sciences. The responses and changes also enhance the robustness of the models and give rise to a framework for a never ending cycle of research on similar phenomena. It can be said that these functions allow the models to become powerful tools for analyzing the dispersion and incidence of variants in a wide range of phenomena by changing critical parameters.

Section 2: Neural Networks

Introduction to Neural Networks in NLP

Let me take a clue what factors motivate the professionals adopt AI based neural networks in NLP. During my research in AI networks applied to natural language processing, I mainly focused on two tasks: categorizing written text produced by AI systems as one of the partner tasks – for example, and selection of orthographic forms for phonological inputs as well. Such tasks have been chosen in order to demonstrate what area can be easily enhanced and flexible with the help of neural networks in particular which in this case include tasks such as writing, classification, and sequence of events. I also want to emphasize the fact that there are a few research manuscripts dealing with the issues of how neural networks can be implemented in textual data hierarchy system so I did my best to be one of them.

Part 1: AI Writing Sample Classification

In fulfilling the first part of the assignment, I shifted my focus towards the development of two models that can identify an AI generated work from a human written one. Analysis section2_data1.csv comprised of the writing samples and their respective binary variable labels was therefore the data source for this analysis. Writing this task involved converting the text data submitted to the ML models into submissions for machine learning, in this particular case testing two models, naive bayes classifier and neural network aiming to determine the better option of the two.

Methodology

I began the multiple steps needed for preparing the NYSE text data by applying the traditional bag-of-words technique. The way it works is that, it translates the text into dimension attributes by counting how many words a given text has, in other words transforming text into a vector that contains the number of entries and how many words. With regard to this change, I had to use CountVectorizer that is a part of sklearn library. This was relevant as deep learning makes use of data that is in an unordered text format so that it can be organized and supplied to the models.

Once I preprocessed the data, I split the data into two parts which were training and testing data in the ratio of 80:20. This partition was essential as it enabled the assessment of the model's ability to generalize since the portion of the dataset was used to train the models and the rest was left for the model to evaluate without having training . In a similar way it is possible to return to the situation with random_state = 42, thereby getting the same results again.

Naive Bayes Model:

The first model that I constructed was a Naive bayes What I used was MultinomialNB which is from the sklearn suite. The architecture of this model – the naive bayes algorithm – and also the ability to work with high dimensional vectors explains why this model can be used in appropriate contexts which involve text. Thus, as a naïve bayes model, I set it and trained with the training data and then tested the effectiveness of the training on the testing dataset. The conclusions of the evaluated model are as high as 79.99% accuracy and affirm the possibility of telling apart another text written by AI from a text written by a human. This performance was taken as the standard to which to compare the measured performance of the neural network model.

Neural Network Model:

Subsequently, I implemented a feed forward neural network as below using Keras. The primary architecture of the network network included an input layer, a single hidden layer with 64 ReLU neurons, and a final sigmoid layer for binary classification. Compiled with Adam optimizer and binary cross-entropy loss function and trained for 10 iterations having the batch size of 32.

As for the performance in terms of the proposed neural network, it was quite satisfactory, as the accuracy achieved on the test images was 96%. Such high performance revealed that features from words were able to be captured sufficiently enough by the neural network while the Naive Bayes model assumed features to be independent. The result proved that deep learning was more efficient than the traditional machine learning algorithms for NLP tasks, especially in the presence of big data.

Results and Discussion

The analysis of the Naive Bayes and the neural network models showed that the performance of the latter is much higher. On the other hand, applying the Naive Bayes model gave a performance that was slightly above 80%. The accuracy achieved by the neural network was 96 percent which was a big improvement over the Naive Model of even higher than 80percent. This improvement, in particular, suggested the applicability of the neural network in finding higher order patterns in the data particularly interactions of many features and not simple ones.

The findings assured that the neural networks have yielded better outcomes when it comes to illustrating the performance of deep learning for NLP, especially for the comprehension of more complex data. However, the general goal of creating neural networks is to be able to express more information, however, it is worth noting that these types of networks are also more expensive and will add more costs in fine tuning such parameters as learning rate, structure of the network and so on. That is why there is a challenge of optimizing model accuracy along with the computational time required to find an optimal suitable model for practical applications.

Part 2: Orthography Prediction

The second assignment, included embedding each phonological representation of a word into a prediction model that aimed at the orthographic form of that word. This translation: translation task underscored that neural networks can do quite difficult simultaneous multipoint transformations of ordered input and output with a certain logic which further supports such use of neural networks in NLP tasks.

Methodology

The orthodata dataset consists of English words presented in their orthographic and phonological transcriptions. The dataset is encased in the csv file format and has been given the title orthodata.csv. In this experiment, the purpose was more on how a certain word can be spelled given its phonological representation provided by the computer. For this purpose, the specific technique that I used was character level encoding where each character is assigned a numeric value based on the database output from LabelEncoder.

Data Preparation:

To start with, both the phonological and orthographic forms were taken out of the dataset and then LabelEncoder was used for encoding the forms. At this stage, the characters were substituted with an integer which, in comparison, made it simpler to analyze the sequences of characters that were being analyzed. The sequences were then padded with a certain amount to make the total length of the sequences equal in order for the sequences to be batch trained during training of the model.

Embedding Model:

The architecture of the model incorporated an embedding layer along with Long Short Term Memory layer and the output dense layer which used Soft Max function as an activation. One of the layers turned a set of sequences of input vectors into dense vectors which were intended to establish that there is a semantic relation of characters on the sequence and the given order. These vectors were movement vectors that used memory to detail the people that had made up previous characters the LSTM layer which works this way has shifted through these vectors appropriately. The dense layer produced a probability distribution of characters at each position and give the sequence of characters which was considered the most orthographically likely to be actually presenting.

I went for 10 epochs with batch size of 32 and the model accuracy that I achieved on the test set stayed around 89.47% on the average. This high accuracy meant that the model was good for example as a model for capturing the relationship of the phonological and the orthographic forms especially for tasks that involved phonological to orthographic transformations to be done.

Results and Discussion

the embedding model demonstrated that it is capable of acquiring high level sequence maps as a function of its phonological inputs for the orthographic form of words. The sliding window and dictionary search recorded a commendable 89.47% as accuracy in generalizing various interpretive forms of words which was a reasonably good performance in phonological-to-orthographic conversion.

To demonstrate the usefulness of using the model to provide specific estimates, five random words from the test set depicted in table 1 and their actual orthographic form are compared to each other. These estimates turned out to be very close to the reality since the actual words through a completely similar criterion model were present in the text. This success strengthened the rationale of employing sequence models I E LSTMs for tasks involving sequences as it gives due regard to context and order of the sequence.

This achievement demonstrated that there exists such transformation that could be performed by the neural network embedding model. The researchers confirmed that LSTMs were especially beneficial in retaining dependencies within the time steps as they were utilized in various language tasks. This opportunity creates endless possibilities of utilizing neural networks on different kinds of NLP such as translation, speech detection and many others.

Conclusion

To perform these tasks, I got to understand the versatility of Neural Networks within the NLP domain. The latter was showcased in the classification task that aimed at demostrating the efficiency of real neural network models in contexts where Naive Bayes models could not work and the former was illustrated in the orthography prediction task stressing the capability of sequence models to change from one type of linguistic sign to another. These models opened many exciting possibilities of employing neural networks to solve different tasks in NLP and provide quite reasonable approaches for those problems. The accomplishment of these tasks expands my comprehension concerning neural networks and helps me to improve more skills and knowledge so that I would be able to tackle such issues in the nearest future.