Multi-modal Deep Learning Performance Prediction

Alex Ahn, Dylan Lewis,
Jonathan Liu, Jered McInerney

RELATED WORK

Learning **Deep Representations** of data has shown to be effective in boosting accuracy on both classification and regression tasks. In addition, by leveraging multiple sources of data, it is possible to learn **hidden correlations** between the data sets. Previous work on Multi-Modal Deep Learning has been done by Andrew Ng using the modalities of speech and video. **Restricted Boltzman Machines** (RBM) have also shown success in learning the deep representations on untagged data. We used MLB game outcomes to predict the next outcome with data from different sources including statistics, social media, and news articles.

MODEL

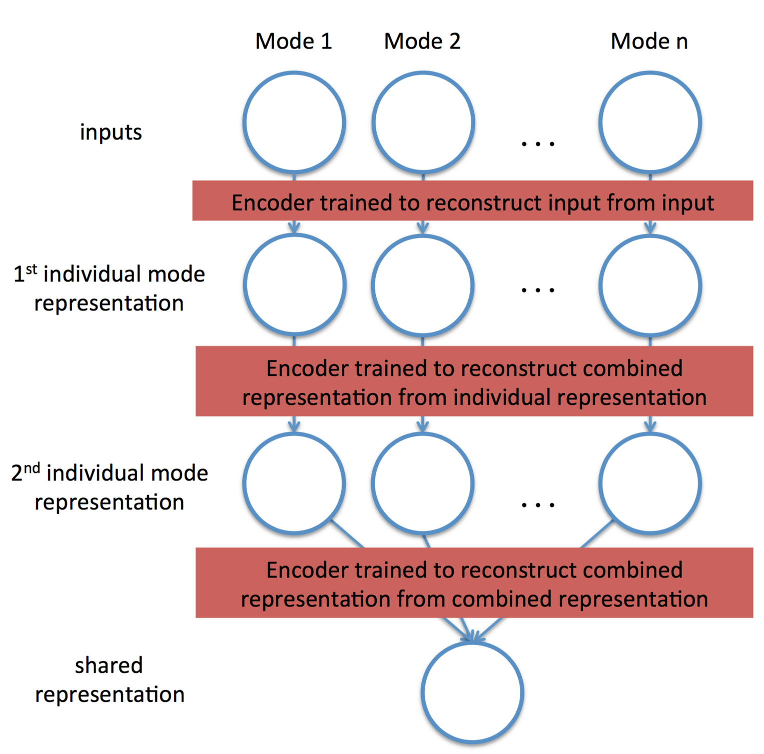


Figure 1. Multimodal Shared Representation Learning

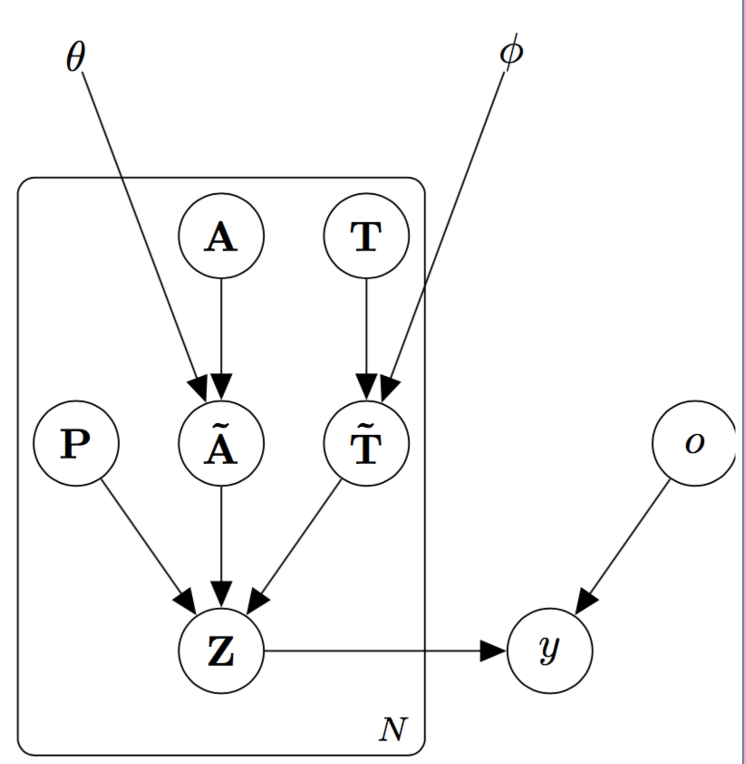
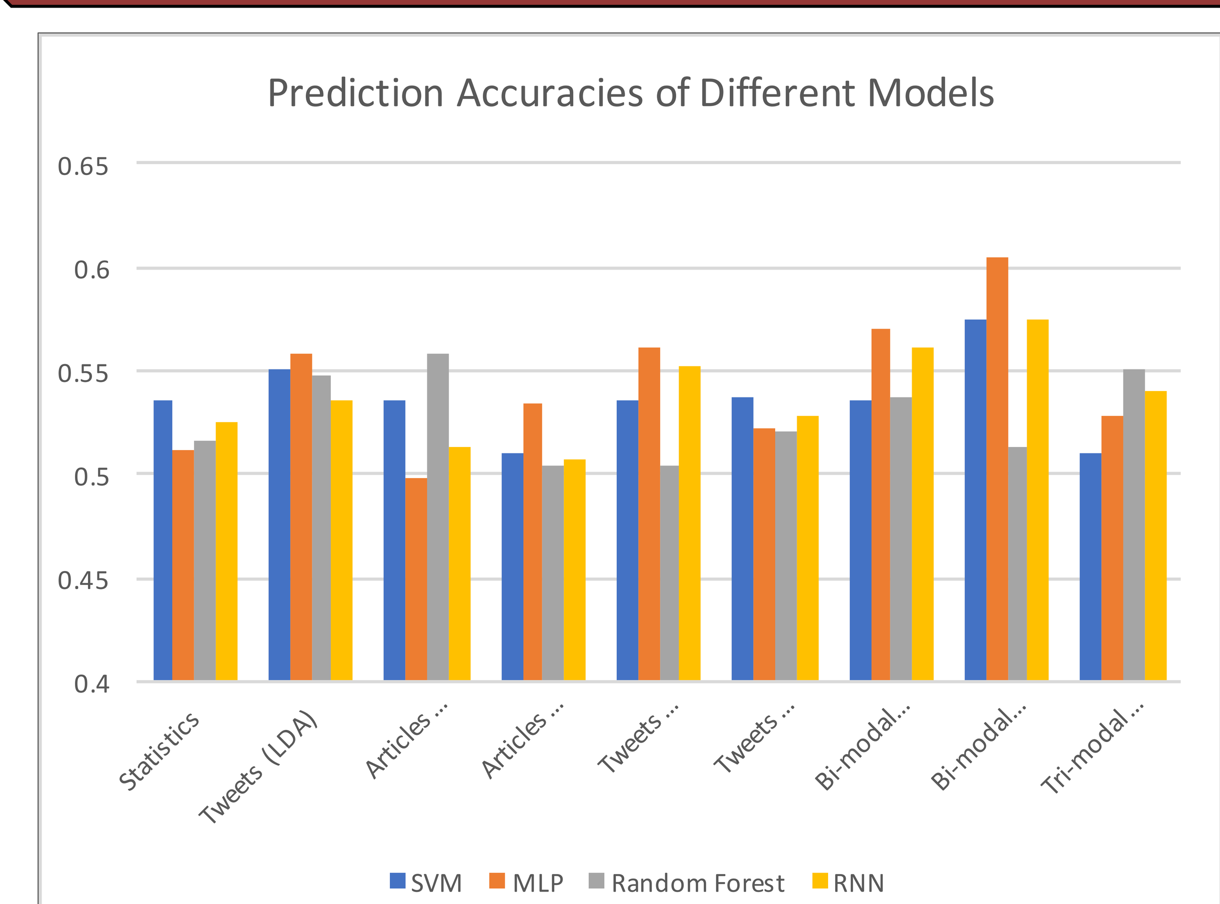


Figure 2. Plate Notation of the Model

FIGURES



OBJECTIVE

1.Extract data from multiple sources with different modalities (Twitter, Baseball Statistics, News Articles)
2.Preprocess the textual data through topic modeling and word embedding.
3.Learn shared representations of the multi modal data using deep learning techniques.
4.Predict the next game outcome using observed data from multiple sources.

METHODS

1. Select the domain of problem which we want to make performance prediction: Orioles games from MLB.
2. Collect data from 3 different sources using keywords (MLB records, Twitter, News Articles)
3. Preprocess the text data using LDA and Word Embedding
4. Perform evaluation on each model with different input modalities: (a) single-modal (b) bi-modal (c) tri-modal
6. For each step, revise the dependencies of each random variables and perform evaluation iteratively. (I.e. one-hot-encoding of categorical variables, splitting latent/observed variables further by time-frame, and accounting for the temporal structure)
7. Construct individualized representations of of multi modal data using deep learning with restricted Boltzmann machines.
8. Use a deep auto-encoder to turn these representations into a shared representation, and run the previous simple models on these new shared representation.
9. Alter the restrictive Boltzmann machines and auto-encoder to append the win/loss label to the output when training encoders, and run the resulting representations on the previous models.
10. Try altering number of layers in encoders to find the best model.

RESULTS

	Statistics	Tweets (LDA)	Articles (LDA)	Articles (Embedding)	Tweets + Articles	Tweets + Articles + Statistics	Bi-modal Shared Rep	Bi-modal Shared Rep with labels	Tri-modal Shared Rep
SVM	0.536	0.551	0.536	0.511	0.536	**0.537**	0.536	0.575	0.511
MLP	0.512	**0.559**	0.498	**0.535**	**0.561**	0.523	**0.571**	**0.605**	0.529
Random Forest	0.517	0.548	**0.559**	0.504	0.504	0.521	0.537	0.513	**0.551**
RNN	**0.525**	0.536	0.514	0.507	0.553	0.528	0.561	0.575	0.541

ANALYSIS

1. We first began with using SVM, Multi Layer Perceptron, and Random Forest just on the previous statistics (single modality). We didn't find an significant signal in when attempting inference with these techniques, we obtained an accuracy of barely above 0.51
2. We then used the same techniques on the other two modalities of Tweets and Articles, and found a more significant signal, the lowest being 0.551 when using and SVM on Tweets, and the highest being 0.587 when using MLP on Tweets.
3. Performing multi-modal deep learning for shared representation of comprehensive data, we were able to reduce the dimensionality at the same time incorporating dependencies from different sources /types of data.
4. Using the representation encoder, we were able to overcome missing data by learning a shared representation with one or two modalities.
5. Unfortunately, the multimodal model did not perform as well as the simple models, so we attempted to improve this model with an intuition that the using labels in the auto-encoder might help us learn representations that are more helpful in predicting win/loss.

REFERENCES

[1] Ngiam, Jiquan, et al. "Multimodal deep learning." *Proceedings of the 28th international conference on machine learning (ICML-11)*. 2011.
[2] Wang, Weiran, et al. "On Deep Multi-View Representation Learning." *ICML*. 2015.