Knowledge Integration Techniques for Sequence to Sequence and Time Series Models

Introduction

With the increase in the amount of data that is available in today's world, using a Deep Learning Neural Network has been made easier. We can use this to build very good complex models that are pattern based. However if we want to go to the next level, we need to look into the integration of knowledge base systems and Deep learning models, these can provide us with great models that perform well and are much more reliable than those that are not integrated. Today we will be giving a brief summary of some of the work that has been done for Time Series and Sequence to Sequence models with integrations to Knowledge base systems.

Key Terms

- Knowledge Based Systems (KBS): Frameworks that consolidate human aptitude for settling on choices are information-based frameworks (KBS). Customarily an information-based framework comprises of an information base which is information reasonably gathered and sorted out by human specialists in different fields, a derivation motor - that depends on the information base for dynamic, a working memory to deal with tasks. The derivation motor can be rule-based, case-based, and so on
- Deep Neural Network: Is a technique that relies on massive amounts of data to find statistical patterns and non-linear relationship. It then uses these patterns to make inferences.
    - Recurrent Neural Network (RNN): A class of a neural network that deals with predication of temporal sequences. This consist of architectures such as Long Short-term Memory (LSTM), and Gated Recurrent Units (GRU) that are used for time series modeling.
- Sequence to Sequence Models: Translates a fixed length input to a fixed length output. It is a family of machine learning approaches used for language processing. Its applications are for language translation, image captioning, conversational models, and text summarization.

Knowledge Integration Techniques

Deep Learning techniques and Knowledge based systems can work together to fill the gaps that are presented from each of the techniques. Deep Learning can improve in areas such as interpretability, and reasoning, while Knowledge Based Systems can improve upon the using statistical patterns to better areas that they are unable to identify.

Time-Series Models

The author goes over five different papers tackling areas such as analysis of popularity of online content, Demand forecasting within a supply chain, a technique for Short-term Load forecasting, a stock trend prediction model, and finally a technique to transform a process of converting a signal in terms of a mother wavelength.

Predicting the popularity of online content with knowledge-enhanced neural networks.

Within this paper a Knowledge base data is spoken to us as an implanting vector that is found out by limiting the loss of triples and then catching the likeness of these triples. The strategy proposes to upgrade the expectation and an LSTM is utilized for anticipating the ubiquity over the long run.

An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain.

In this paper they propose an integration strategy that is inspired by boosting the ensemble model in which decisions are made as a weighted combination of results. The Deep Learning Neural Network and ensemble integration present a better understanding of the supply chain and presents a much more reliable method of forecasting for the future.

Short-term load forecasting by using a combined method of convolutional neural networks and fuzzy time series.

Here we focus on Short-Term Load Forecasting and the integration of fuzzy time series with a convolution neural network. They use a Fuzzy model to convert the time series data into images, then each channel of images corresponds to one of the variables in the data. This allows us to represent the problem with binary answers, strengthening the reliability and accuracy of the model.

Knowledge-driven stock trend prediction and explanation via temporal convolutional network

A common use case for machine learning is to make a stock trend prediction. Within this paper The author uses the structured knowledge in the form of knowledge graph and integrates it into a neural network. They use a Temporal Convolution Network to help predict the stock trend.

Wavelet neural network embedded expert system used in short-term load forecasting

Using a knowledge base we can improve the performance by converting the wavelength into different aspects. The wavelet neural organizations probably won't catch certain elements that are area explicit, in which case the procedure influences on a bunch of rules set up by specialists. In view of the verifiable information, the rationale equation from master experience might be depicted as:

$$forecastV(t) - firstV(t) + D \text{ x } firstV(t)$$

where firstV (t) is the first load value predicted by the neural network at time t; D is the decision making factor at that moment; forecast (t) is the outcome obtained by the expert system predict at time t.

Sequence-to-Sequence Models

Within a Sequence to sequence model the author goes over multiple different use-cases such as automatic speech recognition, an technique for cold fusion, cell control fusion, and an integration of arbitrary knowledge sources for machine translation. He also goes over a technique for iterative distillation, and finally machine comprehension.

Learn spelling from teachers: Transferring knowledge from language models to sequence-to-sequence speech recognition

Within this paper the authors show how to use sequence prediction for automatic speech recognition. A pre-trained model is the prior knowledge and then is integrated with neural network. The language model will not only predict the expected K hard labels but also provide an exhaustive list of soft labels which is used as prior knowledge

Cold fusion: Training seq2seq models together with language models

Cold Fusion uses a sequence to sequence model fused with external language model through a parametric gated function.

$$h_t^{\mathrm{LM}} = \mathrm{DNN}\left(\ell_t^{\mathrm{LM}}\right)$$

$$g_t = \sigma\left(W\left[s_t; h_t^{\mathrm{LM}}\right] + b\right)$$

$$s_t^{\mathrm{CF}} = \left[s_t; g_t \circ h_t^{\mathrm{LM}}\right]$$

$$r_t^{\mathrm{CF}} = \mathrm{DNN}\left(s_t^{\mathrm{CF}}\right)$$

$$\hat{P}\left(y_t | x, y_{<t}\right) = \mathrm{softmax}\left(r_t^{\mathrm{CF}}\right)$$

This can summarize the training process where the model's probability can be determined.

Language model integration based on memory control for sequence to sequence speech recognition

This presents a technique cell control fusion that is similar to but is still different in the aspects of not just fusing the gated outputs of hidden states of the external language model. This will improve the model with a Knowledge base and a neural network combined.

Conclusion

Over time we reach a point where either a knowledge base or a statistical model are not enough to predict or make a good enough model. Combining both of these systems together can help us improve our models and provide us with better accuracy. Having a knowledge

system work with statistical patterns can make a much larger difference in the way the functions perform and the results that are generated.

References

- https://arxiv.org/pdf/2008.05972.pdf
- https://en.wikipedia.org/wiki/Knowledge-based_systems
- https://en.wikipedia.org/wiki/Long_short-term_memory
- https://en.wikipedia.org/wiki/Seq2seq