Deep Probabilistic Models - Exercises

Wilker Aziz
ILLC @ UvA

Outline

1. Recognising Textual Entailment
2. Predicting Publication Date
3. Forecasting Voting Intentions

Recognising Textual Entailment

We have gathered a collection of pairs of textual passages and annotated them for logical entailment, that is, whether a piece of text (the premise) entails or not the other (the hypothesis). Data of this kind has been used to train binary classifiers used to recognise cases of valid entailment relations for tasks such as information retrieval, paraphrasing, and question answering (Dagan et al., 2005).

Examples

Text (or premise): *The two suspects belong to the 30th Street gang, which became embroiled in one of the most noto- rious recent crimes in Mexico: a shootout at the Guadalajara airport in May, 1993, that killed Cardinal Juan Jesus Posadas Ocampo and six others.*

Hypothesis: *Cardinal Juan Jesus Posadas Ocampo died in 1993*

Entailment: True

Recognising Textual Entailment

We have gathered a collection of pairs of textual passages and annotated them for logical entailment, that is, whether a piece of text (the premise) entails or not the other (the hypothesis). Data of this kind has been used to train binary classifiers used to recognise cases of valid entailment relations for tasks such as information retrieval, paraphrasing, and question answering (Dagan et al., 2005).

Examples

Text (or premise): *American Airlines began laying off hundreds of flight attendants on Tuesday, after a federal judge turned aside a union's bid to block the job losses.*

Hypothesis: *American Airlines will recall hundreds of flight attendants as it steps up the number of flights it operates.*

Entailment: False

Exercise

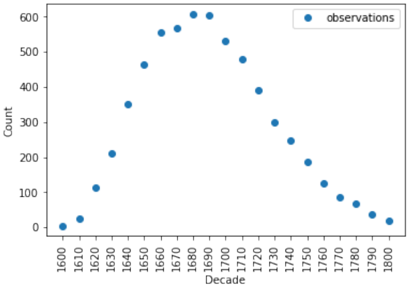
Design a probabilistic model of the data and use NNs to parameterise the likelihood. This model will be used in a question-answering pipeline.

1. You can choose whether to model conditionally or jointly, but motivate your choice.
2. State your choice of distributions clearly.
3. Write down the likelihood of a dataset with N observations as a function of the parameters of the distributions you chose.
4. Give one concrete parameterisation of the model: you may use abstract NN blocks (e.g., embedding layer, FFNN, RNN), but you should explain its inputs and outputs. In particular make sure that your NN maps predictors to the correct parameter space.
5. Propose a concrete algorithm to classify a given pair of passages with a trained model.

Outline

Predicting Publication Date

We have discovered a massive collection of historical documents, most likely written between the years 1600 and 1800. A group of dedicated historians worked hard to annotate a sample of such documents with the decade they were likely written (amazing job!). They managed to annotate 10% of the collection and luckily the sample covers every decade in the interval under consideration.



Exercise

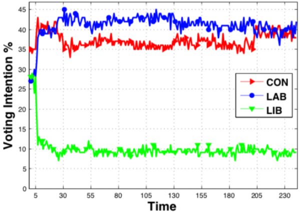
Exercise: Design a probabilistic model of the data and use NNs to parameterise the likelihood. This model will be used to annotate the remaining of the collection.

1. You can choose whether to model conditionally or jointly, but motivate your choice.
2. State your choice of distributions clearly.
3. Write down the likelihood of a dataset with N observations as a function of the parameters of the distributions you chose.
4. Give one concrete parameterisation of the model: you may use abstract NN blocks (e.g., embedding layer, FFNN, RNN), but you should explain its inputs and outputs. In particular make sure that your NN maps predictors to the correct parameter space.
5. Propose a concrete algorithm to predict the decade in which a given document was likely written.

Outline

Forecasting Voting Intentions

Lampos et al. (2013) gathered 240 voting intention polls in the period from April 2010 and February 2012 in the UK. They also collected tweets about politics from user accounts based in the UK. Using the available metadata they organised the tweets on a timeline covering the polling period.



(a) 240 voting intention polls for the 3 major parties in the UK (April 2010 to February 2012)

Forecasting Voting Intentions - Data

- Tweets and voting intentions arranged in a time line.
- Assume time is discretised (say in days), and assume the number of tweets T per unit of time is fixed.
- Also assume that for every K units of time you have access to new polling data (in the form of percentage points) for 3 political parties.

Exercise

Exercise: Design a temporal probabilistic model of the data and use NNs to parameterise the likelihood. This model will be used to forecast voting intentions using a stream of tweets about politics.

1. You can choose whether to model conditionally or jointly, but motivate your choice.
2. State your choice of distributions clearly.
3. Write down the likelihood of a dataset with N observations as a function of the parameters of the distributions you chose.
4. Give one concrete parameterisation of the model: you may use abstract NN blocks (e.g., embedding layer, FFNN, RNN), but you should explain its inputs and outputs. In particular make sure that your NN maps predictors to the correct parameter space.
5. Propose a concrete algorithm to predict voting intentions over 1 to 10 units of time into the future (you may assume availability of tweets if you chose to model conditionally).

References I

Ido Dagan, Oren Glickman, and Bernardo Magnini. The PASCAL recognising textual entailment challenge. In *Machine learning challenges workshop*, pages 177–190, 2005. tex.organization: Springer.

Vasileios Lampos, Daniel Preoțiuc-Pietro, and Trevor Cohn. A user-centric model of voting intention from Social Media. In *Proceedings of the 51st annual meeting of the association for computational linguistics (volume 1: Long papers)*, pages 993–1003, Sofia, Bulgaria, August 2013. Association for Computational Linguistics. URL https://www.aclweb.org/anthology/P13-1098.