Your title here

Student 1 **Student 2** **Student 3**

Feel free to add more sections but those listed here are strongly recommended.

1 Introduction

Natural language is You can keep this short. Ideally you introduce the task already in a way that highlights the difficulties your method will tackle.

2 Methodology

The idea of our approach follows that of successes in previous work such as [5] and [1]. More specifically our approach consists of training a binary logistic regression classifier based on an augmented data set which we create ourselves. This augmentation is based on ideas that have been successfully used in the past (e.g. [4]). As features for our classifier we consider a collection of feature extractors which produce a vector of features for any 5 sentence story.

2.1 Data augmentation

The augmented data set is composed of labeled 5 sentence stories. We create it by combining the original data set and a negative data set. Here the original dataset gets assigned a positive label and the negative set a negative label. We compose the negative dataset by replacing the ending of each 5 sentence story by a randomly selected ending, as done by [4]. We use a replication factor of 1, i.e. for every sample of the original training set we add a single negative sample.

2.2 Feature extraction

In order to fit a logistic regression model as described above we use features obtained from feature extractors. A feature extractor is simply a model which takes in a 5 sentence story and outputs a feature or vector of features. Feature extractors are discussed extensively in section 3. To obtain the final set of features we simply combine features from all extractors for each sample.

3 Model

The core components of our approach can be seen in figure 1. As described in section 2, the C_i are the feature extractors and indicates a mapping:

$$C : \mathbb{S}^5 \rightarrow \mathbb{R}^l \tag{1}$$

That is, it's a function of the set of all 5 sentence stories ($\mathbb{S}^5$) to a real vector.

4 Training

What is your objective? How do you optimize it?

5 Experiments

This **must** at least include the accuracy of your method on the validation set.

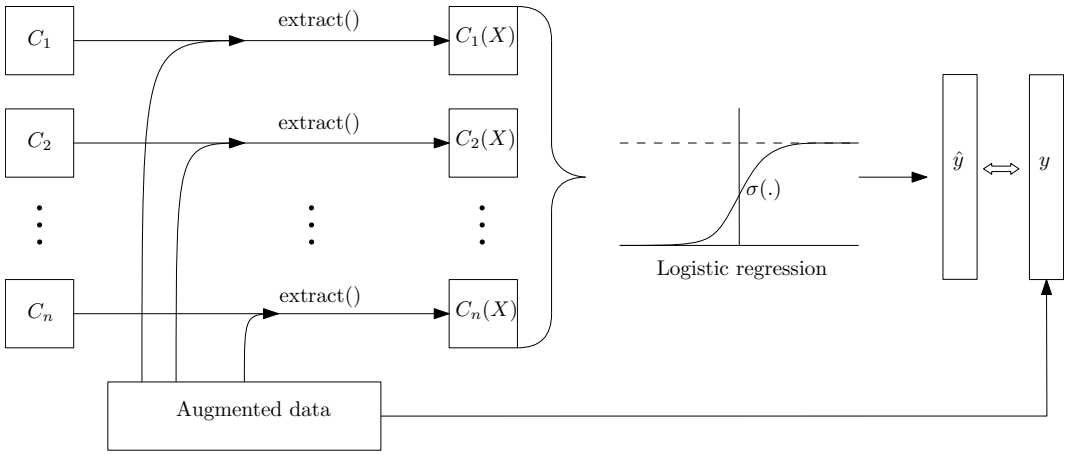

Figure 1: Main components of the global approach.

6 Conclusion

You can keep this short, too. [0]

References

[1] Snigdha Chaturvedi, Haoruo Peng, and Dan Roth. Story comprehension for predicting what happens next. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1603–1614. Association for Computational Linguistics, 2017.

[2] Nasrin Mostafazadeh, Nathanael Chambers, Xiaodong He, Devi Parikh, Dhruv Batra, Lucy Vanderwende, Pushmeet Kohli, and James F. Allen. A corpus and evaluation framework for deeper understanding of commonsense stories. *CoRR*, abs/1604.01696, 2016.

[3] Nasrin Mostafazadeh, Michael Roth, Annie Louis, Nathanael Chambers, and James F. Allen. Lsdsem 2017 shared task: The story cloze test. 2017.

[4] Melissa Roemmele, Sosuke Kobayashi, Naoya Inoue, and Andrew Gordon. An rnn-based binary classifier for the story cloze test. In *Proceedings of the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-level Semantics*, pages 74–80. Association for Computational Linguistics, 2017.

[5] Roy Schwartz, Maarten Sap, Ioannis Konstas, Leila Zilles, Yejin Choi, and Noah A Smith. Story cloze task: Uw nlp system. In *EACL Workshop LSD Sem*, pages 52–55, 2017.