

Master Thesis Proposal

Deep Learning of Humor from Gary Larson's Cartoons

Robert Fischer BSc.

Advisor: Dr. Horst Eidenberger, Assoc. Prof.

May 06, 2018

1 Problem Definition

The main goal of this diploma thesis is to research whether computers are capable of understanding human humor by learning Gary Larson's Cartoons. This is done by applying several Deep Learning techniques.

Existing approaches mainly focused on humor in written form, for example: Joke Recognition and Pun Generation. With the advent of Deep Learning and convolutional neural networks (=CNN) in recent years, images may now be taken into consideration as well and similarly recurrent neural networks have been improving the state-of-the-art in several natural language processing disciplines. Research of humor using deep learning is still in its infancy. Computational humor is a relatively new field of computer science and deep learning as well, has had its breakthrough also not too long ago in XXX.

Convolutional neural networks are usually applied to real world pictures, the question remains whether they are also capable of analyzing cartoons as well and if there are ways on how to improve CNNs for this type of images.

Another goal is to evaluate the resulting neural networks. Hidden inside them, there might be some interesting insights about human humor, especially since convolutional neural networks are inspired by the human brain. Maybe each neural network layer will have its own dedicated function, or there will be a cow-detection neuron inside the network? The question, if a deep neural network is able to detect whether a cartoon contains a certain type of animal is of interest. Convolutional neural networks are state-of-the-art for real world image classification, the question remains if this is also the case for cartoony images.

2 Expected Results

The main outcome of this thesis is a machine learning pipeline which is able to classify a cartoon by funniness. This pipeline consists of a pre-processing phase, training phase and testing phase. In the pre-processing phase the available cartoons are pre-processed in such a way that the deep learning networks can be properly trained. Additionally the ground truth has to be annotated by humans. In the training phase, a model is generated using a subset of the ground truth, which also generalizes to previously unseen cartoons. Therefore several neural network configurations will be evaluated and the goal is to find the best for the given ground truth. In the test phase the results are evaluated using measures like F-Score and accuracy.

Additionally one finds answers to the following questions:

- Are convolutional neural networks able to detect certain types of animals in Gary Larson's cartoons?
- Will the final neural network reveal insights of human humor?
- To what degree is deep learning able to understand human humor by learning Gary Larson's cartoons?

3 Methodology and Approach

3.1 Pre-Processing

First the ground truth will be created: A data set of several thousand cartoons exists, but these cartoons need to be preprocessed in such a way that a neural network can be trained on this data set: The punchlines have to be extracted, the pictures need to be normalized and bad examples filtered. Additionally manual labelling of cartoons is necessary: How funny a cartoon is, as well as what animals are visible have to be annotated.

3.2 Humor Classification

Given this data set, based on previous classifications, the goal is to predict how funny a previously unseen cartoon is. This is done by training multiple machine learning classifiers. Important to note is the fact, that humor is subjective. So the generated model might only be applicable for individuals.

For the visual component of a cartoon a convolutional neural network will be used, as they have been applied very successfully in various image classification tasks. Several configurations will be evaluated. If and how deep neural networks have to be adapted to work with cartoons will be researched during the thesis.

A punchline analysis will be implemented using recurrent neural networks and other natural language processing techniques. Prime candidate are Long Short-Term Memory networks, but other techniques will be considered as well. One key problem of recurrent neural networks is that, they tend to overfit to the training data very easily.

Finally both neural networks are merged together using another, possibly more traditional, classifier.

3.3 Animal Classification

A similar approach will be implemented for animal classification of Gary Larson's cartoons. The question remains how to handle comics properly using deep neural networks, additionally some data augmenting techniques might be implemented to avoid overfitting.

3.4 Deep Neural Network Evaluation

The final neural networks will be evaluated. The idea is that, there might be some

4 State of the Art

TODO

5 Relation to Master Study

TODO