Neural networks with statistical learning

Sören Schleibaum, TU Clausthal 26. Oktober 2016

1 Ideas for a project

Questions

- Which competitions could use a neural network?
- @Aditya: What's your favorite?

Kaggle competitions

I've marked the ones, which I think are interesting, bold.

- **Outbrain Click Prediction**
  - Deadline: 18th of January 2017
  - Data:
  - Task:
  - **Outbrain** is a content discovery platform
  - Predict which pieces of content its global base of users are likely to click on in a way that the recommendation for undiscovered stories will satisfy the individual tastes of users better.
  - Recommendation algorithm
- **Leaf Classification**
  - Deadline: 28th of February 2017
  - Data: Binary leaf images + extracted features
  - Task: Build a classifier.
  - Automating plant recognition might have many applications, for instance species popoulation tracking and preservation.
  - Use binary leaf images and extracted features to accurately identify 99 species of plants.
- House Prices: Advances Regression Techniques
  - Deadline: 1st of March 2017
  - Data:
  - Task:

  - 79 explanatory variables describing residential homes in Iowa
  - Try to predict the final price of each home.
- **Dogs vs. Cats Redux: Kernels Edition**
  - Deadline: 2nd of March 2017
  - Data:
  - Task:
  - Dog vs. Cat classification problem
  - I think it's about image recognition.
- Allstate Claims Severity
  - Deadline: 12th of December 2016
  - Data:
  - Task:
  - Developing automated methods of predicting cost and severity of claims.
  - Create an algorithm which accurately predicts claims severity.
- Transfer Learning on Stack Exchange Tags
  - Deadline: 25th of March 2017
  - Data: Titles, text, ant tags on Stack Exchange questions from six different sites
  - Task: Predict tags of unseen physics questions.
  - Might involve training an algorithm on a corpus
- Ghouls, Goblins, and Ghosts... Boo!
  - Deadline: 1st of December 2017
  - Data:
  - Task: Classification
  - Gradient boosting machines
- **Facial Keypoints Detection**
  - Deadline: 31st of December 2016
  - Data:
  - Task: Predict keypoint positions on face images.
  - Has some helping introduction and tutorials.
- **Can you pair products with people**
  - Deadline: 21st of December 2016
  - Data:
  - Task: Predict which products the customers will use based on past behavior.
  - Personalized product recommendation

2 Implementation Steps

- Normalize all of your images, both for training and testing, to have the same resolution.
- Use gray-scale images, so each pixel would give you just one number.
- Image Thresholding to process image based on intensity.
- Extracting features (e.g., edges) from the image and then using the network on those features. This incorporates prior knowledge.
- Use each pixel value as one input to your network. For instance, if you have images of size 16x16 pixels, your network would have 16*16 = 256 input neurons. The first neuron would see the value of the pixel at (0,0), the second at (0,1), and so on.

  The first neuron would see the value of the pixel at (0,0), the second at (0,1), and so on. Basically you put the image values into one vector and feed this vector into the network.