Automatic Metadata Extraction with Conditional Random Fields

Joseph Boyd

École Polytechnique Fédérale de Lausanne

joseph.boyd@epfl.ch

April 3, 2015

Outline

Project Objectives

- Lorem ipsum dolor sit amet, consectetur adipiscing elit
- Aliquam blandit faucibus nisi, sit amet dapibus enim tempus eu
- Nulla commodo, erat quis gravida posuere, elit lacus lobortis est, quis porttitor odio mauris at libero
- Nam cursus est eget velit posuere pellentesque
- Vestibulum faucibus velit a augue condimentum quis convallis nulla gravida

Outline

Logistic Regression

- A logistic regression is used for classifying a data sample into two (binary) or more (multi) categories, thus,

$$\hat{y}_{prediction} = \boldsymbol{\beta}^T \cdot \mathbf{x}_{sample},$$

where $\hat{y}$ is the prediction (represented as a probability), $\mathbf{x} = [x_0, x_1, ..., x_D]^T$ is a data sample, and $\boldsymbol{\beta} = [\beta_0, \beta_1, ..., \beta_D]^T$ is the vector of parameters we must *learn*

- We construct a (maximum log likelihood) cost function in terms of this parameter vector,

$$\mathcal{L}(\boldsymbol{\beta}) = \sum_{n=1}^{N} y_n \boldsymbol{\beta}^T \mathbf{x}_n - \log[1 + \exp(\boldsymbol{\beta}^T \mathbf{x}_n)]$$

Solving a Logistic Regression

- Building a regression model is equivalent to solving a convex optimisation problem (i.e. maximising the cost function)
- We know the form of the model, and we have a set of (training) data
- We want to choose the model parameters for which the error is minimised (think line of best fit)
- We use a numerical method to obtain the global minimum of error, for example, the method of gradient descent:

$$\boldsymbol{\beta}^{k+1} = \boldsymbol{\beta}^{k} - \alpha \nabla \mathcal{L}(\boldsymbol{\beta}^{k})$$

Take home message: we can automatically build mathematical functions for making predictions

Hidden Markov Models (HMMs)

- Lorem ipsum dolor sit amet, consectetur adipiscing elit
- Aliquam blandit faucibus nisi, sit amet dapibus enim tempus eu
- Nulla commodo, erat quis gravida posuere, elit lacus lobortis est, quis porttitor odio mauris at libero
- Nam cursus est eget velit posuere pellentesque
- Vestibulum faucibus velit a augue condimentum quis convallis nulla gravida

Hidden Markov Models (HMMs) - Example

- Lorem ipsum dolor sit amet, consectetur adipiscing elit
- Aliquam blandit faucibus nisi, sit amet dapibus enim tempus eu
- Nulla commodo, erat quis gravida posuere, elit lacus lobortis est, quis porttitor odio mauris at libero
- Nam cursus est eget velit posuere pellentesque
- Vestibulum faucibus velit a augue condimentum quis convallis nulla gravida

Solving Hidden Markov Models

- Solved using dynamic programming techniques
- Aliquam blandit faucibus nisi, sit amet dapibus enim tempus eu

Take home message: once we have the model, we can make predictions for a given input *efficiently*.

Conditional Random Fields (CRFs)

- Lorem ipsum dolor sit amet, consectetur adipiscing elit
- Aliquam blandit faucibus nisi, sit amet dapibus enim tempus eu
- Nulla commodo, erat quis gravida posuere, elit lacus lobortis est, quis porttitor odio mauris at libero
- Nam cursus est eget velit posuere pellentesque
- Vestibulum faucibus velit a augue condimentum quis convallis nulla gravida

Outline

Outline

Outline