TRDDC @ FIRE 2013

System for Classification of Legal Propositions

Forum for Information Retrieval Evaluation

December, 2013

Nitin Ramrakhiyani Sachin Pawar

Tata Research Development and Design Centre

Tata Consultancy Services Ltd.

Introduction

Introduction

- Submitted as a participation in Task 2(b) of the Information Access in the Legal Domain Track.

Introduction

- Submitted as a participation in Task 2(b) of the Information Access in the Legal Domain Track.
- Goal: To classify Legal Propositions into respective classes of legal importance like arguments, facts, etc.

Introduction

- Submitted as a participation in Task 2(b) of the Information Access in the Legal Domain Track.
- Goal: To classify Legal Propositions into respective classes of legal importance like arguments, facts, etc.
- Data

Introduction

- Submitted as a participation in Task 2(b) of the Information Access in the Legal Domain Track.
- Goal: To classify Legal Propositions into respective classes of legal importance like arguments, facts, etc.
- Data
  - Training
    - 9 Classes namely FI, FE, I, A, LR, SS, SP, SO and R
    - 712 Propositions (with class information)

Introduction

- Submitted as a participation in Task 2(b) of the Information Access in the Legal Domain Track.
- Goal: To classify Legal Propositions into respective classes of legal importance like arguments, facts, etc.
- Data
  - Training
    - 9 Classes namely FI, FE, I, A, LR, SS, SP, SO and R
    - 712 Propositions (with class information)
  - Testing
    - 390 Plain Propositions

Methodology

[1] B. Hachey and C. Grover, "Sentence classification experiments for legal text summarisation," in *Proc. 17th Annual Conference on Legal Knowledge and Information Systems (Jurix-2004)*, 2004, pp. 29–38

[2] A. L. Berger, V. J. D. Pietra, and S. A. D. Pietra, "A maximum entropy approach to natural language processing," *Computational linguistics*, vol. 22, no. 1, pp. 39–71, 1996

Methodology

- Initial Attempt: Using Naive Bayes for classification with bag-of-words features

[1] B. Hachey and C. Grover, "Sentence classification experiments for legal text summarisation," in *Proc. 17th Annual Conference on Legal Knowledge and Information Systems (Jurix-2004)*, 2004, pp. 29–38

[2] A. L. Berger, V. J. D. Pietra, and S. A. D. Pietra, "A maximum entropy approach to natural language processing," *Computational linguistics*, vol. 22, no. 1, pp. 39–71, 1996

Methodology

- Initial Attempt: Using Naive Bayes for classification with bag-of-words features

- Devised a set of features to represent each proposition.
  - We use a set of 16 features in our system.
  - Some features based on work in Hachey and Grover[1]

[1] B. Hachey and C. Grover, "Sentence classification experiments for legal text summarisation," in *Proc. 17th Annual Conference on Legal Knowledge and Information Systems (Jurix-2004)*, 2004, pp. 29–38

[2] A. L. Berger, V. J. D. Pietra, and S. A. D. Pietra, "A maximum entropy approach to natural language processing," *Computational linguistics*, vol. 22, no. 1, pp. 39–71, 1996

Methodology

- Initial Attempt: Using Naive Bayes for classification with bag-of-words features
- Devised a set of features to represent each proposition.
  - We use a set of 16 features in our system.
  - Some features based on work in Hachey and Grover[1]
- Trained a suitable classifier using training data.
  - We used a Multi-class Maximum Entropy[2] classifier for this task.

[1] B. Hachey and C. Grover, "Sentence classification experiments for legal text summarisation," in *Proc. 17th Annual Conference on Legal Knowledge and Information Systems (Jurix-2004)*, 2004, pp. 29–38

[2] A. L. Berger, V. J. D. Pietra, and S. A. D. Pietra, "A maximum entropy approach to natural language processing," *Computational linguistics*, vol. 22, no. 1, pp. 39–71, 1996

Methodology

- Initial Attempt: Using Naive Bayes for classification with bag-of-words features
- Devised a set of features to represent each proposition.
  - We use a set of 16 features in our system.
  - Some features based on work in Hachey and Grover[1]
- Trained a suitable classifier using training data.
  - We used a Multi-class Maximum Entropy[2] classifier for this task.
- Testing the classifier model on the test set.

---

[1] B. Hachey and C. Grover, "Sentence classification experiments for legal text summarisation," in *Proc. 17th Annual Conference on Legal Knowledge and Information Systems (Jurix-2004)*, 2004, pp. 29–38

[2] A. L. Berger, V. J. D. Pietra, and S. A. D. Pietra, "A maximum entropy approach to natural language processing," *Computational linguistics*, vol. 22, no. 1, pp. 39–71, 1996

Feature Set

Feature Set

- Bag of Words: Unordered set of proposition tokens post stop-word removal

Feature Set

- Bag of Words: Unordered set of proposition tokens post stop-word removal
- Presence of Named Entities namely PERSON, ORGANIZATION and LOCATION

Feature Set

- Bag of Words: Unordered set of proposition tokens post stop-word removal
- Presence of Named Entities namely PERSON, ORGANIZATION and LOCATION
- Proposition Length (Encoded as SHORT, MEDIUM and LONG)

Feature Set

- Bag of Words: Unordered set of proposition tokens post stop-word removal
- Presence of Named Entities namely PERSON, ORGANIZATION and LOCATION
- Proposition Length (Encoded as SHORT, MEDIUM and LONG)
- Presence of Month and Year patterns

Feature Set

- Bag of Words: Unordered set of proposition tokens post stop-word removal
- Presence of Named Entities namely PERSON, ORGANIZATION and LOCATION
- Proposition Length (Encoded as SHORT, MEDIUM and LONG)
- Presence of Month and Year patterns
- Presence of a monetary value like Rs 32,000.00

Feature Set

- Bag of Words: Unordered set of proposition tokens post stop-word removal
- Presence of Named Entities namely PERSON, ORGANIZATION and LOCATION
- Proposition Length (Encoded as SHORT, MEDIUM and LONG)
- Presence of Month and Year patterns
- Presence of a monetary value like Rs 32,000.00
- Proposition Position i.e. Location in the actual article.

Feature Set

Feature Set

- Presence of FIRST PERSON, SECOND PERSON, THIRD PERSON pronouns

Feature Set

- Presence of FIRST PERSON, SECOND PERSON, THIRD PERSON pronouns
- Presence of at least one past and future tense verb

Feature Set

- Presence of FIRST PERSON, SECOND PERSON, THIRD PERSON pronouns
- Presence of at least one past and future tense verb
- Presence of Legal domain words like “section”, “hereinafter referred to”, “court”, “Act”, “Rule”, “Order”

Feature Set

- Presence of FIRST PERSON, SECOND PERSON, THIRD PERSON pronouns
- Presence of at least one past and future tense verb
- Presence of Legal domain words like “section”, “hereinafter referred to”, “court”, “Act”, “Rule”, “Order”
- Presence of quotes (“)

Feature Set

- Presence of FIRST PERSON, SECOND PERSON, THIRD PERSON pronouns
- Presence of at least one past and future tense verb
- Presence of Legal domain words like “section”, “hereinafter referred to”, “court”, “Act”, “Rule”, “Order”
- Presence of quotes (“)
- Point Wise Mutual Information (PMI) feature $PMI(w, t)$
  $Score(w,\ t) = \frac{No.\ of\ propositions\ of\ type\ t\ containing\ w}{No.\ of\ propositions\ of\ type\ t \times Frequency\ of\ w}$

Evaluation

Evaluation

- Tried with different classifiers and different feature combinations. Best performance with the above set of features and Maxent classifier

Evaluation

- Tried with different classifiers and different feature combinations. Best performance with the above set of features and Maxent classifier
- Performance Measurement: Micro-averaged accuracy obtained in 10-fold cross validation of training data.

Evaluation

- Tried with different classifiers and different feature combinations. Best performance with the above set of features and Maxent classifier
- Performance Measurement: Micro-averaged accuracy obtained in 10-fold cross validation of training data.
- Recorded accuracy: 62.78% $\longrightarrow$ 52% (Testing)

Evaluation

- Tried with different classifiers and different feature combinations. Best performance with the above set of features and Maxent classifier
- Performance Measurement: Micro-averaged accuracy obtained in 10-fold cross validation of training data.
- Recorded accuracy: 62.78% $\longrightarrow$ 52% (Testing)
- Best performing class: FI
  F-measure: 0.72 $\longrightarrow$ 0.85 (Testing)

Evaluation

Evaluation

- Propositions of type I, A, SP, FE and SO present scantily in the training data.

Evaluation

- Propositions of type I, A, SP, FE and SO present scantily in the training data.
- The FI and R confusion
  - FI Mostly concentrated towards the upper part
  - R Mostly concentrated towards the lower part
  - Major discerning features: Position and Pronouns

Evaluation

- Propositions of type I, A, SP, FE and SO present scantily in the training data.
- The FI and R confusion
  - FI Mostly concentrated towards the upper part
  - R Mostly concentrated towards the lower part
  - Major discerning features: Position and Pronouns
- Still space for more features like TF-IDF of terms, Parse/POS oriented features.

Evaluation

- Propositions of type I, A, SP, FE and SO present scantily in the training data.
- The FI and R confusion
  - FI Mostly concentrated towards the upper part
  - R Mostly concentrated towards the lower part
  - Major discerning features: Position and Pronouns
- Still space for more features like TF-IDF of terms, Parse/POS oriented features.
- Need for more domain expertise to understand errors

Questions ?