Relation extraction methods for biomedical literature

Bui, Q.C.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

UvA-DARE is a service provided by the library of the University of Amsterdam (http://dare.uva.nl)
# Contents

1. Introduction  
1.1 Motivation  
1.2 Research questions  
1.3 Outline of the thesis  

2. Relation extraction methods for biomedical text  
2.1 Introduction  
2.1.1 Text preprocessing  
2.1.2 Named entity recognition  
2.1.3 Parsing  
2.1.4 Relation extraction  
2.1.5 Evaluation metrics  
2.2 Natural Language Processing  
2.2.1 Sentence splitting  
2.2.2 Lexical processing  
2.2.3 Syntactic processing  
2.3 Machine learning  
2.3.1 SVM methods for relation extraction  
2.3.2 Kernel methods  
2.4 Biomedical corpora  
2.4.1 GENIA corpus  
2.4.2 PPI corpora  
2.4.3 GENIA events corpus  
2.5 Relation extraction methods  
2.5.1 Co-occurrence approaches  
2.5.2 Pattern-based approaches  
2.5.3 Rule-based approaches  
2.5.4 Machine learning-based approaches  
2.5.5 Performance comparison of existing approaches to relation extraction  
2.6 Conclusion  

3. Extracting causal relations on HIV drug resistance from literature  
3.1 Introduction  
3.2 System and methods  
3.2.1 Text retrieval  
3.2.2 Text preprocessing  
3.2.3 Relation extraction  
3.2.4 Relation combination  
3.3 Results and Discussion  
3.3.1 Datasets  
3.3.2 Relation extraction performance  
3.3.3 Relation combination performance  
3.4 Conclusions  

4. A hybrid approach to extract protein-protein interactions  
4.1 Introduction
## Contents

4.2 System and methods

  4.2.1 Text preprocessing
  4.2.2 Extracting candidate PPI pairs

4.3 Results and Discussion

  4.3.1 Data sets
  4.3.2 Evaluation methods
  4.3.3 Performance of PPI extraction algorithm
  4.3.4 Single corpus evaluation
  4.3.5 Cross-corpora evaluation
  4.3.6 Performance time

4.4 Conclusions

5. A robust approach to extract biomedical events from literature

  5.1 Introduction

5.2 System and methods

  5.2.1 Structured representation
  5.2.2 Learning rules
  5.2.3 Rule combination
  5.2.4 Event extraction
  5.2.5 Text preprocessing

5.3 Results and discussion

  5.3.1 Datasets
  5.3.2 Evaluation settings
  5.3.3 Event extraction
  5.3.4 Performance time
  5.3.5 Performance analysis

5.4 Conclusion

6. Discussion and Conclusion

  6.1 The use of syntactic information for different relation extraction tasks
  6.2 The role of machine learning to relation extraction tasks
  6.3 The performance time of relation extraction systems
  6.4 Contribution
  6.5 Future work

References

Summary

Samenvatting

List of publications

Acknowledgements