

## Context specific text mining for annotating protein interactions with experimental evidence

[Login \(/login\)](#)

- [IUPUI ScholarWorks Repository](#)
- →
- [School of Informatics and Computing](#)
- →
- [Informatics Theses and Dissertations](#)
- →
- [Informatics Graduate Theses and PhD Dissertations](#)
- →
- [View Item](#)

## Context specific text mining for annotating protein interactions with experimental evidence

[Pandit, Yogesh](#)



Name: [ypandit\\_bioinfo\\_t ...](#)  
Size: 10.76Mb  
Format: PDF  
Description: Yogesh Pandit's ...  
[View/Open](#)

Permanent Link: <http://hdl.handle.net/1805/3809>  
Date: 2014-01-03  
Committee Chair: [Palakal, Mathew J.](#)  
Committee: Liu, Yunlong  
Members: Liu, Xiaowen  
Degree: M.S.  
Degree Year: 2013  
Department: School of Informatics  
Grantor: Indiana University  
LC Subjects: [Protein-protein interactions -- Databases -- Research](#); [Data mining -- Analysis](#); [Systems biology -- Methodology](#); [Bioinformatics -- Information resources -- Research](#); [Machine learning -- Methodology -- Research](#); [Computational intelligence -- Research -- Methodology -- Analysis](#); [Information storage and retrieval systems](#); [Natural language processing \(Computer science\) -- Research](#); [Biology -- Data processing -- Research](#); [Artificial intelligence -- Medical applications -- Research](#)

### Abstract:

Proteins are the building blocks in a biological system. They interact with other proteins to make unique biological phenomenon. Protein-protein interactions play a valuable role in understanding the molecular mechanisms occurring in

any biological system. Protein interaction databases are a rich source on protein interaction related information. They gather large amounts of information from published literature to enrich their data. Expert curators put in most of these efforts manually. The amount of accessible and publicly available literature is growing very rapidly. Manual annotation is a time consuming process. And with the rate at which available information is growing, it cannot be dealt with only manual curation. There need to be tools to process this huge amounts of data to bring out valuable gist than can help curators proceed faster. In case of extracting protein-protein interaction evidences from literature, just a mere mention of a certain protein by look-up approaches cannot help validate the interaction. Supporting protein interaction information with experimental evidence can help this cause. In this study, we are applying machine learning based classification techniques to classify and given protein interaction related document into an interaction detection method. We use biological attributes and experimental factors, different combination of which define any particular interaction detection method. Then using predicted detection methods, proteins identified using named entity recognition techniques and decomposing the parts-of-speech composition we search for sentences with experimental evidence for a protein-protein interaction. We report an accuracy of 75.1% with a F-score of 47.6% on a dataset containing 2035 training documents and 300 test documents.

## Description:

Indiana University-Purdue University Indianapolis (IUPUI)

## This item appears in the following Collection(s)

- [Informatics Graduate Theses and PhD Dissertations \(/handle/1805/303\)](/handle/1805/303)
- [Informatics School Theses and Dissertations \(/handle/1805/954\)](/handle/1805/954)



[Show Statistical Information \(#\)](#)

## My Account

- [Login](#)
- [Register](#)

## Statistics

- [Most Popular Items](#)
- [Statistics by Country](#)
- [Most Popular Authors](#)

[About Us \(/page/about\)](/page/about) | [Contact Us \(/contact\)](/contact) | [Send Feedback \(/feedback\)](/feedback)

[\\_\(/htmlmap\)](#)

## FULLFILLING *the* PROMISE

[Privacy Notice \(http://ulib.iupui.edu/privacy\\_notice\)](http://ulib.iupui.edu/privacy_notice)



Copyright (<http://www.iu.edu/copyright/index.shtml>) ©2015

The Trustees of Indiana University (<http://www.iu.edu/>),

Copyright Complaints (<http://www.iu.edu/copyright/complaints.shtml>)