Negation detection in medical texts

Maria Chiara Martinis[1,2][0000−0002−4868−7767], Chiara Zucco[1,2][0000−0003−0048−0457], and Mario Cannataro[1,2][0000−0003−1502−2387]

Data Analytics Research Center, University "Magna Græcia", 88100 Catanzaro,Italy
Department of Medical and Surgical Sciences, University "Magna Græcia", 88100 Catanzaro, Italy
{martinis, chiara.zucco, cannataro}@unicz.it

Abstract. Negation detection refers to the automatic identification of linguistic expression that convey negation within a textual content. In medical and biomedical context, the negation detection plays a pivotal role in understanding clinical documentation and extracting meaningful insights. In this paper, we survey 16 articles published from 2005 to 2023 and focusing on negation detection within medical domain. Our evaluation framework encompass both methodological aspects and application-oriented considerations. Specifically, we discuss the used approaches, the emplyed methodology, the specific tasks addressed, the target language of textual analysis, and the evaluation metrics used. On the application front, for each reviewed study, we delineate the medical domains under investigation (e.g., cardiology, oncology), the types of data analyzed, and the availability of datasets. The majority of reviewed works are conducted in English, with a prevalence of machine learning and deep learning approaches, and classic classification evaluation metrics. Application domains exhibit heterogeneity, with a slight predominance in oncology, and diverse data sources including EHRs, abstracts, scientific papers, and web-derived information (e.g., Wikipedia or blog entries). Throughout this review, we will identify limitations and gaps in this research area, as well as examine the benefits it could bring to the scientific community and the methods currently employed.

Keywords: NLP· Negation detection· medical domain

1 Introduction

Negation detection (ND) is a critical element in the field of Natural Language Processing (NLP) and has significant implications for a wide range of applications. This process involves recognizing instances in the text where a statement is negated or its meaning is reversed, proving crucial for understanding conveyed information in Morante et al.[10].

Its importance lies in the fact that negation can completely change the interpretation of a statement, often adding complexity to text comprehension. For example, distinguishing between "The patient is not experiencing a fever" and

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

"The patient has a fever" has substantial implications in decision-making in the healthcare sector.

Negation detection requires in-depth text analysis, including lexical indicators, syntactic patterns, and contextual clues.

Recent developments in the field of Natural Language Processing (NLP), particularly the use of machine learning models such as deep neural networks and transformer-based architectures, have sparked growing interest, promising new perspectives to overcome these challenges.

The goal of this article is to provide an overview of the state of the art in "negation detection," specifically identifying the latest methodologies and emerging technologies, emphasizing the specific linguistic challenges influencing this research field.

We will analyze the practical applications of negation detection in key sectors, such as medicine, highlighting how accurate negation identification can significantly improve result accuracy and relevance.

Through this investigation, we aim to offer an in-depth overview of current frontiers in negation detection research, identify ongoing challenges, and propose possible directions for future developments.

One of the early methods in the literature, explained by Chapman et al. in [1], introduces the negation identification method called NegEx. This approach relies on a set of predefined rules to detect the presence of negations in texts.

In this study, the authors verify, using their algorithm with preprocessed sentences, whether the conclusions and diseases indexed from discharge summaries are negated or affirmed by the dictating physician.

NegEx, despite its usefulness in identifying negations in very complex sentences, has significant limitations. Incorrectly assigning negations to analyzed sentences, particularly within patient treatment diagnoses, could lead to an inaccurate diagnosis analysis.

For this reason, Mehrabi et al.[9], developed a new negation algorithm called DEEPEN, specifically designed to address the false positives of NegEx. The system was developed and tested using electronic health record (EHR) data from Indiana University (IU) and subsequently evaluated on the Mayo Clinic dataset to assess its generalizability.

By incorporating the dependencies of NegEx, DEEPEN can reduce the incorrect negations that NegEx assigns.

The present work aims to review 16 articles on negation detection within medical domain published from 2005 to 2023 to provide an overview of existing methodologies related to the identification of "negation' across different medical-related written texts.

Our evaluation framework encompasses methodological aspects as well as application-oriented considerations. More in details, for each reviewed work, we analyze the proposed approaches, the methods employed, the specific tasks addressed (i.e. negation recognition and/or negation scope resolution), the target language of textual analysis, and evaluation metrics utilized. Regarding appli-

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

cations, we examine the medical domains under investigation (e.g., cardiology, oncology), the source of data analyzed, and dataset availability.

The rest of the paper is organized as follows: Section 2 provides a brief overview of the main approaches to negation detection; Section 3 describes the main approaches to negation detection. Finally, Section 4 concludes the document.

2 Basic definition and open problems

Negation detection and negation scope negation resolution (NSR) are two fundamental methods in the field of Natural Language Processing (NLP). Both of these concepts are useful for analyzing and interpreting natural language accurately, providing tools for various applications in the field of Natural Language Processing.

Negation detection refers to identifying, within a text, expressions indicating the negation of a specific statement or concept.

The capability is fundamental for correctly understanding the meaning of the text and interpreting the information accurately as in Morante et al.[10]. On the other hand, scope negation resolution involves identifying the expression or concept negated within a sentence or a broader context.

In Dalloux et al.[2], the parts of the text affected or involved by a specific negation are identified. The determination described is crucial for accurately interpreting the meaning of a sentence and may vary depending on the linguistic context and the structure of the sentence itself.

The detection of negation in the field of Natural Language Processing (NLP) is of fundamental importance, especially when dealing with medical texts. The presence of negations can significantly impact understanding of the meaning of individual sentences or entire texts within the medical context.

In addressing this complexity, both basic and advanced approaches, as well as specific considerations for negation detection in a scientific context, are outlined. An important challenge in negation detection is represented by the linguistic diversity found in scientific texts. Languages used in scientific contexts often employ technical and specialized languages with complex syntactic structures, making the accurate identification of negations challenging.

The use of negated expressions through elaborate syntactic constructs requires an advanced approach to ensure precise detection.

Beyond the difficulties faced in identifying negation in scientific documents, a critical aspect involves resolving scope negation. This term pertains to recognizing or expanding the negation within a singular sentence or a more extensive context. A precise interpretation hinges on a thorough understanding of how negation impacts the message's structure.

Ultimately, this process leads to a heightened understanding of how negation operates in the message, pinpointing the scope in which it is employed.

The detection of negation in Natural Language Processing (NLP) can present several challenges, including:

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

1. Linguistic Diversity:
   Languages show variations in grammatical structure and how negation is expressed, posing a challenge to establishing universal rules or models;

2. Technical Language:
   When dealing with scientific or technical texts, the presence of specialized terminology and intricate syntactic structures can impede the precise identification of negation;

3. Ambiguity:
   Negation might be conveyed in an unclear or implicit manner, demanding an understanding of the contextual nuances for precise detection;

4. Availability of Datasets for Specific Languages or Domains:
   The presence of datasets suitable for training models may be limited, particularly for languages that are less common or specific professional domains.

To address these challenges, various types of approaches are highlighted, such as the utilization of rules-based techniques, machine learning, and deep learning to provide more sophisticated solutions. Among the rules-based approaches, the use of linguistic rules emerges as a fundamental starting point. By defining keywords associated with negation, such as "non," "neither," and "none," it's possible to structure an identification mechanism that handles cases where negation is expressed through complex syntactic constructs [?].

In parallel, lexical analysis proves useful in creating a list of negative terms, employing specific techniques to detect them in the text. The application of Part-of-Speech (POS) tagging allows for the examination of the grammatical structure of sentences, identifying typical patterns associated with negation.

Advanced approaches go further, adopting techniques from Machine Learning and Deep Learning. Machine learning involves training a classifier based on a dataset annotated with negative information. Models like recurrent neural networks (RNN) or transformer neural networks (BERT) demonstrate high performance in NLP tasks.

Embedding word vectors, representing words as vectors, leverages semantic similarity to identify words associated with negation. Specific considerations for scientific articles include attention to technical terminology, handling the complex structures often present in scientific publications, and collecting or creating specific datasets for model training and evaluation.

Adapting the model to the language and scientific context, along with the use of appropriate metrics such as precision, recall, and F1-score, contributes to ensuring an accurate evaluation of the system's performance. When experimenting with different approaches, the goal is to find the most suitable solution for the specific context of scientific analysis.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

3 Main approaches to negation detection

In this section, we analyze the methodologies employed to detect negation, with a particular focus on the medical context.

The analysis involves various strategies, such as rule-based approaches (3.1), machine learning models (3.2), and deep learning techniques (3.3), to address challenges, as shown in Table 1.

Table 1. Example of Approaches and Methodologies for Negation Detection in Different Languages.

Authors	Approaches	Year
Elkin et al.[3]	Rules-based	2005
Zamaraeva et al.[17]	Rules-based	2018
Hammami [6]	Rules-based	2021
Mutalik et al.[12]	Machine learning	2001
Chapman et al.[1]	Machine learning	2001
Huang et al.[7]	Machine Learning-Ruled based	2007
Vincze et al.[16]	Machine Learing	2008
Savova et al.[14]	Machine learning	2010
Funkner et al.[5]	Machine Learning	2014
Mukherjee et al.[11]	Machine Learning	2017
Sun et al.[15]	Machine Learning	2021
Mehrabi [9]	Deep Learning	2015
Dalloux [2]	Deep Learning	2019
Daan de Jong [8]	Deep Learning	2021
Van Es et al.[4]	Deep Learning	2023

3.1 Rule-Based Approaches

Rule-based approaches rely on predefined rules and linguistic models to recognize negation expressions in this context. The following articles belong to this category. Hammami et al. [6] introduce a novel rule-based approach, employing natural language processing (NLP) techniques for classifying Italian pathological reports based on the ICD-OM coding scheme.

This approach facilitates the identification and categorization of morphological content in Italian pathological reports by analyzing negation patterns. The method involves selecting two negation categories during training, assigning identified negation terms to these categories, and merging negation terms with negated terms into single bigrams based on their positions. The final classification algorithm achieved a noteworthy micro-F1 score of 98.14% across 9594 pathological reports in the test dataset.

The study proposed by Elkin et al.[3], aims to compare the effectiveness of an automated system in assigning negation to clinical concepts within compositional

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

expressions with manually assigned negation. The dataset comprises 41 clinical documents (medical evaluations) analyzed through the Mayo Vocabulary Server Parsing Engine, utilizing SNOMED-CT™ for conceptual coverage.

Validation of the identification of concepts and textual cues related to negation is conducted through a review by a medical terminologist. The results of recall were 97.2% and a specificity of 98.8% in negation assignment, indicating high precision compared to manual evaluation.

The Zamaraeva et al. in [17] focus on optimizing feature extraction in pathology reports, with particular attention to precise negation detection. The authors propose a targeted approach to accurately recognize and delineate the context of negation within medical texts.

This improvement in negation detection contributes to a more accurate extraction of relevant information from pathological reports, enhancing the overall quality of the extracted features. The document introduces new methodologies for negation detection and discusses the results obtained, highlighting the potential impact on medical practice.

3.2 Machine Learning Approaches

Mukherjee et al.[11] propose a generic parser called NegAIT (Negation Assessment and Inspection Tool) is being developed and thoroughly examined to annotate the presence of negations in a text. This parser takes text as input and identifies occurrences of morphological negations, negative phrases, and double negations.

It is implemented in an integrated Java-Scala environment, using a rule-based approach through the Open Domain Informer (ODIN) event extraction framework. The steps to recognize negation include text tokenization, sentence splitting, word stemming and analysis with the Stanford parser in [**?**].

Manually crafted rules combining regular expressions and lexicons are applied to identify various types of negation. Due to the initial lexicon generating many false positives, additional negative words from other lexicons have been incorporated.

The work in [5], focuses on the development of a negation detection module for the analysis of unstructured clinical documents in the Russian language, employing a machine learning-based approach. The module has been trained and tested using anonymized electronic health records of patients with acute coronary syndrome. The notable effectiveness of the module is particularly evident in predicting the potential need for surgical interventions for patients affected by acute coronary syndrome.

The adopted methodology involves annotating clinical documents for specific pathologies, normalizing annotated texts, and utilizing a gradient-boosting classifier for training and optimization. Relevant outcomes emerge in refining predictive models related to surgical interventions, based on text characteristics, while the impact on outcome prediction models is less pronounced.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

This negation detection module represents a crucial component within a broader application for clinical document analysis, enhancing prediction accuracy, and contributing significantly to advanced practices in the clinical context.

Sun et al.[15] discusses the participation of the MedAI system in SemEval-2021 Task 10, focusing on domain adaptation for negation detection. The authors introduce a new method called "Negation-aware Pre-training," designed to enhance the model's ability to detect negations in source-free contexts.

The study explores the challenges associated with negation detection in the medical context and outlines the methodology employed by the MedAI system. This approach may involve utilizing pre-training techniques specifically crafted to handle negations and adapting the model to the unique characteristics of the medical domain. The document also presents experimental results, highlighting the effectiveness of the proposed approach in improving negation detection performance in medical text, even without access to labeled source data.

Naldi et al.[13], present an evaluation method based on a test set to assess the effectiveness of a sentiment analysis tool in identifying negations within a medical sentence. They examined a basic test set containing over thirty manually labeled sentences. However, they observed that widely used sentiment analysis packages continue to exhibit inefficiency in handling negative sentences, primarily due to their test set-based nature. This approach forces the algorithm to deal with highly critical sentences, whose polarity may be challenging for the algorithm itself to comprehend.

A potential solution could involve adopting machine learning-based approaches. However, this requires a substantial amount of labeled data for training, which makes challenging the training of such approaches.

The negation annotator in cTAKES utilizes the NegEx algorithm in [**?**], based on models, to identify words and phrases indicative of negation near mentions of named entities. Similarly, the status annotator employs a similar approach to identify relevant words and phrases indicating the status of a named entity.

For negation detection, each discovered named entity is associated with one of the dictionary's semantic types and includes attributes such as the associated text span ('span'), the associated terminology/ontology code ('concept'), the negation of the named entity ('negation'), and the status associated with the named entity ('status').

The status value is set to "possible" for future events, while allergies to a specific drug are managed by setting the negation attribute of that drug to "is negated." Non-patient experiences are marked as "family history of" if applicable.

cTAKES is distributed with highly performant machine learning modules and models and operates on the Apache Unstructured Information Management Architecture (UIMA) URL http://incubator.apache.org/uima/. Vincze et al.[16], analyze a corpus annotated for negations, speculations, and their linguistic extensions, emphasizing that detecting signals of negation and uncertainty is more straightforward in clinical documents due to the abundance of keywords.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

The study highlights the importance of detecting negations and scopes in the biomedical context, with over 10% of sentences containing modifiers that significantly influence semantic meaning. The BioScope corpus is designed to facilitate the development of automatic systems for detecting negations and scopes, addressing challenges related to keyword identification and linguistic scope determination.

Mutalik et al.in [12], test the hypothesis that the detection of negated concepts in dictated medical documents can be effectively achieved using tools designed for the analysis of formal languages. They developed a program called Negfinder, incorporating a lexical scanner and a parser based on a subset of context-free grammars.

This was analyzed on a diverse set of 40 medical documents to recognize negation patterns in the text. The parser's performance was evaluated on two test sets: one visually inspected for false positives and false negatives, and another independently examined by both a human observer and Negfinder, ultimately comparing the results and achieving a sensitivity of 95.7%.

Jamai et al.in [7], a novel hybrid method is outlined, which integrates the use of regular expression matching with grammatical analysis. This approach aims to overcome the previously mentioned constraint in automatically detecting negations within clinical radiology reports. Jamai et al. in [14] aim to develop

and evaluate an open-source natural language processing system designed to extract information from free-text electronic health records. The system, named cTAKES (Clinical Text Analysis and Knowledge Extraction System), is available as open source at http://www.ohnlp.org and extensively leverages existing open-source technologies, including the Unstructured Information Management Architecture framework and the OpenNLP natural language processing toolkit. cTAKES has been specifically trained for the clinical domain and demonstrates a high precision of 94.9%.

3.3 Deep Learning Approaches

Daan de Jong in [8], presents a strategy based on the use of neural networks to resolve the scope of predicted negation cues in two phases. The main objective is to address the challenge of accurately identifying the scope of words or phrases indicating negation.

The first phase focuses on predicting negation cues, while the second phase is dedicated to resolving the associated scope of these cues. In the initial phase, a neural network is trained to predict negation cues in the text. Subsequently, in the second phase, another neural network is employed to resolve the scope of these predicted negation cues. By utilizing a two-step approach, the aim is to enhance accuracy in identifying the meaning or scope of negations within the context of the text.

In this study [2] , Dalloux et al. analyze a natural language processing (NLP) technique, specifically "negation detection", in French biomedical texts. To con-

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

duct this analysis, the authors have first identified the words that express speculation and negation. Subsequently, they identified their contexts, namely, the tokens within the sentences that are influenced by the presence of negation or speculation. This approach was tested in two French datasets, annotated with negation and speculation signals along with their respective scopes.The approach examined utilizes CRF and BiLSTM, achieving precision results of 97.21%.

Van Es et al. in [4] provide a detailed explanation of the implementation of both approaches, discussing the results obtained through a thorough and comparative evaluation. Metrics such as precision, recall, and F1-score are analyzed to assess the performance of the methods in different conditions and clinical contexts. The authors also examine cases where one approach might be preferable over the other, considering the specific challenges present in Dutch clinical texts.

The ultimate goal of the paper is to offer a clear guide on choosing between rule-based and machine learning methods for negation detection in Dutch clinical texts, taking into account the effectiveness and practicality of each approach in specific clinical contexts.

Mehrabi et al. in [9] developed a new negation algorithm called DEEPEN, specifically designed to address the false positives of NegEx. The system was developed and tested using electronic health record (EHR) data from Indiana University (IU) and subsequently evaluated on the Mayo Clinic dataset to assess its generalizability. By incorporating the dependencies of NegEx, DEEPEN can reduce the incorrect negations that NegEx assigns.

4 Discussions and Results

By delving into the challenges and methodologies analyzed in the preceding sections, our objective is to highlight the role of negation detection in ensuring the accurate comprehension of information in medical texts, significantly enhancing the quality of analyses conducted.

As stated before, our evaluation framework encompasses two main aspects: methodological aspects as well as application-oriented insights. More in details, we evaluated the 16 reviewed works methodological aspects based on:

- approach types (rule-based, ML-based, DL-based);
- proposed methodology;
- the specific tasks addressed i.e. Negation Detection or Negation Scope Resolution (NSR);
- the target language od the analysis,
- the evaluation metrics.

Table 2 summarize the findings. When considering the application-oriented aspects, the focus is posed on the application areas such as medical specialties, the source of the analyzed data, and dataset availability.

Among the issues that emerged, significant challenges are identified stemming from the substantial lack of datasets specifically compiled in a suitable technical language to correctly identify negation expressions.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

Table 2. The table shows information about the proposed methodology, the specific task, the target language, and the metrics used to compare the 16 works under review.

Article	Methods	Task	Language	Metrics
Elkin et al. [3]	Rules	ND	English	Precision
Zamarava et al. [17]	Rules	ND	English	Not Available
Hammami et al. [6]	Rules	ND	Italian	F1 score
Mutalik et al. [12]	Regular expressions	ND	English	Not Available
Chapman et al. [1]	Regular expression	ND	English	Specificity
Huang et al. [7]	Hybrid approaches, regualar expression matching with grammatic parsing	ND	English	Specificity
Vincnze et al. [16]	Annotation dataset	ND-NSR	English	Not Available
Savova et al. [14]	Machine Learning	ND	English	Accuracy and F1 score
Funkner et al. [5]	Multiclass classification	ND	Russian	F1 score
Mukherjee et al. [11]	Binary classification	ND	English	Precision and Recall
Sun et al. [15]	Self-supervised Model	ND	English	F1 score
Meharabi [9]	Graph-based and transition-based	ND	English	Precision and Recall
Dalloux [2]	BiLSTM	ND-NSR	French	Precision
Daan de Jong [8]	BiLSTM	NSR	English	F1 score
Van Es et al. [4]	BiLSTM	ND	German	Not Available

This can make the training process of negation detection models difficult, limiting the ability to adequately extract the linguistic nuances present in medical texts.

In Table 3 a detailed representation of the application scopes addressed in the articles and the availability of the training/evaluation dataset is provided. The discussed analysis will be important for identifying emerging trends and assessing the influence of various methodologies in the field of negation detection. Furthermore, the ambiguity of the data and the approaches used could add

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

further complexity to these challenges, requiring a more careful and detailed analysis.

Table 3. The table shows the application domain, source and availability of the dataset

Article	**Application area**	**Source**	**Dataset Availability**
Elkin et al. [3]	Internal Medicine	Electronic Health Records	Not Available
Zamarava et al. [17]	Oncology	Pathological reports	Available
Hammami et al. [6]	Oncology	Pathological reports of oncology	Not Available
Mutalik et al. [12]	Surgery	medical documents	Not Available
Chapman et al. [1]	Not Available	Electronic Health Records	Not Available
Huang et al. [7]	Radiology	Clinical Radiology Reports	Not Available
Vincze et al. [16]	Not Available	Medical reports	Available
Savova et al. [14]	Not Available	Clinical reports	MAYO
Funkner et al. [5]	Cardiology	Electronic Medical Reports related to acute coronary syndrome	Not Available
Mukherjee et al. [11]	Not Available	6 corpora: patient blogs, Cochrane reviews, PubMed abstracts, clinical trial texts, and English Wikipedia articles for different medical topics	Not Available
Sun et al. [15]	Not Available	Clinical notes	i2b2 2010 Challenge Dataset, available under data use agreements
Mehrabi [9]	Not Available	Electronic Health Records	MAYO clinical
Dalloux [2]	Cardiology, Urology, Oncology	French biomedical documents	Available
Daan de Jong [8]	Radiology and Biological research	Radiology clinical reports, full papers, and scientific abstracts in the biological domain	Bioscope corpus
Van Es et al. [4]	Not Available	Clinical corpus	Erasmus Dutch Clinical Corpus

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

Our study has emphasized the importance of negation detection in ensuring accurate comprehension of information in medical texts, thus enhancing the quality of analyses conducted.

One of the issues highlighted in this paper is the limited availability of datasets specifically designed for negation detection, especially in the medical field. This shortage represents a barrier to the advancement of precise models for this technique. As a future work, this could be addressed by creating specific data sets for negation detection with a particular focus on medical domains.

Given this, future efforts could focus not only on generating new datasets but also on adapting and expanding existing linguistic resources. This initiative would facilitate the development of advanced existing models, leading to further growth in research in this field.

Acknowledgement

This work was funded by the Next Generation EU - Italian NRRP, Mission 4, Component 2, Investment 1.5, call for the creation and strengthening of 'Innovation Ecosystems', building 'Territorial R&D Leaders' (Directorial Decree n. 2021/3277) - project Tech4You - Technologies for climate change adaptation and quality of life improvement, n. ECS0000009. This work reflects only the authors' views and opinions, neither the Ministry for University and Research nor the European Commission can be considered responsible for them.

References

1. Chapman, W.W., Bridewell, W., Hanbury, P., Cooper, G.F., Buchanan, B.G.: A simple algorithm for identifying negated findings and diseases in discharge summaries. Journal of biomedical informatics **34**(5), 301–310 (2001)
2. Dalloux, C., Claveau, V., Grabar, N.: Speculation and negation detection in French biomedical corpora. In: Mitkov, R., Angelova, G. (eds.) Proceedings of the International Conference on Recent Advances in Natural Language Processing (RANLP 2019). pp. 223–232. INCOMA Ltd., Varna, Bulgaria (Sep 2019). https://doi.org/10.26615/978-954-452-056-4_026, https://aclanthology.org/R19-1026
3. Elkin, P.L., Brown, S.H., Bauer, B.A., Husser, C.S., Carruth, W., Bergstrom, L.R., Wahner-Roedler, D.L.: A controlled trial of automated classification of negation from clinical notes. BMC medical informatics and decision making **5**, 1–7 (2005)
4. van Es, B., Reteig, L.C., Tan, S.C., Schraagen, M., Hemker, M.M., Arends, S.R., Rios, M.A., Haitjema, S.: Negation detection in dutch clinical texts: an evaluation of rule-based and machine learning methods. BMC bioinformatics **24**(1), 10 (2023)
5. Funkner, A., Balabaeva, K., Kovalchuk, S.: Negation detection for clinical text mining in russian. Studies in health technology and informatics **270**, 342–346 (06 2020). https://doi.org/10.3233/SHTI200179
6. Hammami, L., Paglialonga, A., Pruneri, G., Torresani, M., Sant, M., Bono, C., Caiani, E.G., Baili, P.: Automated classification of cancer morphology from italian pathology reports using natural language processing techniques: A rule-based approach. Journal of Biomedical Informatics

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

116, 103712 (2021). https://doi.org/https://doi.org/10.1016/j.jbi.2021.103712, https://www.sciencedirect.com/science/article/pii/S1532046421000411

7. Huang, Y., Lowe, H.J.: A Novel Hybrid Approach to Automated Negation Detection in Clinical Radiology Reports. Journal of the American Medical Informatics Association **14**(3), 304–311 (05 2007). https://doi.org/10.1197/jamia.M2284, https://doi.org/10.1197/jamia.M2284
8. de Jong, D.: Scope resolution of predicted negation cues: A two-step neural network-based approach. CoRR **abs/2109.07264** (2021), https://arxiv.org/abs/2109.07264
9. Mehrabi, S., Krishnan, A., Sohn, S., Roch, A.M., Schmidt, H., Kesterson, J., Beesley, C., Dexter, P., Schmidt, C.M., Liu, H., et al.: Deepen: A negation detection system for clinical text incorporating dependency relation into negex. Journal of biomedical informatics **54**, 213–219 (2015)
10. Morante, R., Blanco, E.: Recent advances in processing negation. Natural Language Engineering **27**(2), 121–130 (2021). https://doi.org/10.1017/S1351324920000534
11. Mukherjee, P., Leroy, G., Kauchak, D., Rajanarayanan, S., Romero Diaz, D.Y., Yuan, N.P., Pritchard, T.G., Colina, S.: Negait: A new parser for medical text simplification using morphological, sentential and double negation. Journal of Biomedical Informatics **69**, 55–62 (2017). https://doi.org/https://doi.org/10.1016/j.jbi.2017.03.014, https://www.sciencedirect.com/science/article/pii/S1532046417300631
12. Mutalik, P.G., Deshpande, A., Nadkarni, P.M.: Use of general-purpose negation detection to augment concept indexing of medical documents: a quantitative study using the umls. Journal of the American Medical Informatics Association **8**(6), 598–609 (2001)
13. Naldi, M., Petroni, S.: A testset-based method to analyse the negation-detection performance of lexicon-based sentiment analysis tools. Computers **12**(1) (2023). https://doi.org/10.3390/computers12010018, https://www.mdpi.com/2073-431X/12/1/18
14. Savova, G.K., Masanz, J.J., Ogren, P.V., Zheng, J., Sohn, S., Kipper-Schuler, K.C., Chute, C.G.: Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): architecture, component evaluation and applications. Journal of the American Medical Informatics Association **17**(5), 507–513 (09 2010). https://doi.org/10.1136/jamia.2009.001560, https://doi.org/10.1136/jamia.2009.001560
15. Sun, J., Zhang, Q., Wang, Y., Zhang, L.: MedAI at SemEval-2021 task 10: Negation-aware pre-training for source-free negation detection domain adaptation. In: Palmer, A., Schneider, N., Schluter, N., Emerson, G., Herbelot, A., Zhu, X. (eds.) Proceedings of the 15th International Workshop on Semantic Evaluation (SemEval-2021). pp. 1283–1288. Association for Computational Linguistics, Online (Aug 2021). https://doi.org/10.18653/v1/2021.semeval-1.183, https://aclanthology.org/2021.semeval-1.183
16. Vincze, V., Szarvas, G., Farkas, R., Móra, G., Csirik, J.: The bioscope corpus: biomedical texts annotated for uncertainty, negation and their scopes. BMC bioinformatics **9**(11), 1–9 (2008)
17. Zamaraeva, O., Howell, K., Rhine, A.: Improving feature extraction for pathology reports with precise negation scope detection. In: Bender, E.M., Derczynski, L., Isabelle, P. (eds.) Proceedings of the 27th International Conference on Computational Linguistics. pp. 3564–3575. Association for Computational Linguistics, Santa Fe, New Mexico, USA (Aug 2018), https://aclanthology.org/C18-1302

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6

18. Frederic. (2014). Categorization and Construction of Rule Based Systems. Communications in Computer and Information Science. 459. 10.1007/978-3-319-11071-4-18.

ICCS Camera Ready Version 2024
To cite this paper please use the final published version:
DOI: 10.1007/978-3-031-63772-8_6