



# Detecting Contradictions



Halicioğlu Prodigies  
April 27th, 2023



# Solution Overview

How can we use AI to assist with finding contradictions in government policy documents?



## INPUT DATA

### Data Sets

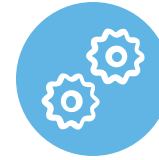
- Our model is evaluated using the DoD Publications & Air Force Publications as one dataset.
- Our model is trained on various datasets that have manually labeled contradictions.
- Basic manipulation is done on the input data to ensure pages are combined into one document.



## DATA MANIPULATION

### Haystack Usability

- Utilized open-source preprocessing libraries already available (Haystack)
- Rids of titles, headers, footers, table of contents, and more.
- Allows for sentences to be found.
- Has ability to "chunk" sentences using a sliding window approach.



## TRANSFORMER USAGE

### MiniLM Transformer

- all-MiniLM-L6-v2 used for embedding of text into a common vector-space for similarity comparison.
- Once text is embedded in chunks, we utilized cosign similarity to score the chunk similarity.
- Similarity scoring used as basis for finding pairs of chunks with similar context.



## CONTRADICTION SCORING

### ROBERTA Contradiction Evaluation

- Robustly Optimized BERT Approach, a variant of BERT (BiDirectional Encoder Representations from Transformers)
- Contextualizes representations of words in a sentence.
- Trained on datasets containing contradictions. Model would be improved if trained on Government data, but not possible without manually creating labeled dataset.

## Just want to see it in action?

Follow along with our notebook, including example results at the bottom!



[https://nbviewer.org/github/mlevitt-deloitte/Policy\\_Recon\\_HDSI/blob/main/DoD%20Contradictions.ipynb](https://nbviewer.org/github/mlevitt-deloitte/Policy_Recon_HDSI/blob/main/DoD%20Contradictions.ipynb)

# Design Considerations

## Our Solution

We endeavored to make applicability a guiding principle during development. Namely, *how easy will it be for the PolicyRecon™ team to **adapt** and **deploy** our solution to fit future business and client needs?*

### ADAPT

#### Open-Source

The full codebase required to run our solution is hosted on GitHub for understandability and mutability.



#### Reusable Pipelines

We've developed our solution as several performant pipeline stages which can be applied as is, modified, isolated, or swapped out as desired.



### DEPLOY

#### Self-Contained

We support a Docker-based development environment which contains all dependencies for plug-and-play accessibility.



#### Scalable & Productionizable

We've deployed to Databricks in the DS Lab — a Deloitte-managed Azure environment which leverages a central data lake, powerful compute and analysis tools, and easy operationalization for client engagements.



# Discovering Similarity

Utilized Haystack to chunk sentences using a sliding window approach. Utilized MiniLM transformer embedding to find chunks with similar context.

## Innovative

It is most common to group chunks of text with paragraphs, but this was not a good method since many legal documents are not oriented in a consistent paragraph structure.

Our solution is a sliding window approach since it is a unique way of grouping text without the need for pre-defined paragraphs.

## Efficient

We could just brute force compare all sentences to each other, but this is very expensive computationally and in terms of time.

With our grouped method, we are only comparing pairs of sentences in similar chunks of text, so this massively reduces the amount of contradiction scores we are calculating.

## Effective

With brute force method many pairs of sentences will be about completely different topics, so even if they have a high contradiction score, it is about the context not actual contradictions.

Sentences regarding similar topics will have the contradictions since they will have a different meaning regarding a concept instead of different meaning overall.

## Important

There is no simple way to find contradictions in text because of the need for the sentences to have similar context as well as description of that context.

Our method finds the general context of a section of text with usage of the chunk of text since context is derived from sections not individual sentences. Then our method finds contradictions from these similar sections with contradictions now measuring difference in description instead of general content.

# Identifying Contradiction Candidates



## Contradiction Model

Leverage a state-of-the-art NLI transformer model



## Candidate Heuristic

Select the top  $k$  most contradictory sentence pairs



## Document Retrieval

Directly link to documents for fast human verification

- Our model is already pre-trained, meaning we can leverage state-of-the-art capabilities without having to spend expensive compute resources for training.
- Even **without a labeled training dataset** we can still use a pre-trained model.
- The output of the model is a selection of the top  $k$  most contradictory sentence pairs as candidates for human analysis.
- This reduces the computational time and manual effort needed to review entire documents and **narrows the scope for a human with proper domain knowledge** to quickly identify contradictions.
- Once results are found, the candidates directly link back to the documents where the contradictory sentences (and their surrounding context) are found.
- This **improves explainability and reduces verification time**.

# Results Overview

## Outcome

- Our model was **successful** in finding sentences that were **potentially contradictory** of each other.
- The model was able to find contradictions from passages discussing the **same context across different documents**.
- With the model we found **different types of contradictions**.
- To run our model on **1124 DoD files** it only took **60-80 minutes** with a single, 16-core CPU machine.

## Considerations

- When there is a contradiction for different documents, they may contradict on a sentence level but referencing different contexts.
- Headings were found to be contradictory of each other.
- Paragraphs may have sentences discussing the same thing slightly differently and considered contradictions.
- When it came to results, it is important to consider we do not have the domain knowledge to be able to determine contradictions.

# Result Examples



## Numerical Contradiction

Title: Defense Manpower Data Center Domain Values for Military Personnel Data Extracts

Requires a **3-year** active-duty service agreement.

----

Title: Defense Manpower Data Center Domain Values for Military Personnel Data Extracts

Requires a **6-year** active-duty service agreement.



## Different Proposing Party

Title: Department Of Defense Privacy Program

SUMMARY: **The Defense Logistics Agency proposes** to alter a system of records notice in its inventory of record systems subject to the Privacy Act of 1974 (5 U.S.C. 552a), as amended.

----

Title: Department Of Defense Privacy Program

SUMMARY: **The Department of the Army is proposing** to amend a system of records notice in its existing inventory of records systems subject to the Privacy Act of 1974, (5 U.S.C. 552a), as amended.



## Contradicting Actions

Title: Visual Information (VI)

This Instruction: a. **Reissues DoD Instruction (DoDI) 5040.02** (Reference (a)) in accordance with the authority in DoDI 5122.05 (Reference (b)) to implement policy and assign responsibilities consistent with Office of Management and Budget Circular A-130

----

Title: Visual Information (VI) Productions

d. **Incorporates and cancels DoDI 5040.09** (Reference (f)).



## Subject Disagreement

Title: Manual of Military Decorations and Awards: DoD Joint Decorations and Awards

5)The **JSCM** for specific achievement may be awarded with the "R" device to distinguish that the award was earned for the direct hands-on employment of a weapon system or other warfighting activities that had a direct and immediate impact on a combat operation or other military operation (i.e., outcome of an engagement or specific effects on a target),

----


Title: Manual of Military Decorations and Awards: DoD Joint Decorations and Awards

4)The **JSAM** for specific achievement may be awarded with the "R" device to distinguish that the award was earned for the direct hands-on employment of a weapon system or other warfighting activities that had a direct and immediate impact on a combat operation or other military operation (i.e., outcome of an engagement or specific effects on a target),

# Future Vision


- Host as a web-app
- Incorporate human feedback to create labeled training set for improvement
- Develop additional data cleaning and filtering to only compare relevant docs
- Could be scaled in the DS Lab or deployed to CfG Workshop so Model pipeline can integrate with other Deloitte NLP assets

## Import Files





Drag Files Here


---


 Filename.pdf

☒

 Filename2.pdf



 Filename3.pdf



Discard

Get Results

## Results

Number of Contradictions: 5

Title: Filename 4 Sentence 1	Title: Filename 6 Sentence 2	<input checked="" type="checkbox"/> <input type="checkbox"/>
Title: Filename 1 Sentence 1	Title: Filename 4 Sentence 2	<input checked="" type="checkbox"/> <input type="checkbox"/>
Title: Filename 3 Sentence 1	Title: Filename 5 Sentence 2	<input checked="" type="checkbox"/> <input type="checkbox"/>

Exit

Export Contradictions



# Solution Links

See below for links to the pre-processor, model, and referenced studies/datasets we used for our solution

---

## GitHub:

[mlevitt-deloitte/Policy\\_Recon\\_HDSI](#)

## Tools Used:

Pre-Trained RoBERTa NLI Model [roberta-large-snli\\_mnli\\_fever\\_anli\\_R1\\_R2\\_R3-nli](#)

Pre-trained MiniLM embedding model [all-MiniLM-L6-v2](#)

Haystack [PreProcessor Node](#)

Deloitte [Data Science Lab](#)

## Referenced Studies:

[Profile Consistency Identification for Open-domain Dialogue Agents \(aclanthology.org\)](#)

[Search for contradiction | Papers With Code](#)

[A logical-based corpus for cross-lingual evaluation | Papers With Code](#)

[Deep learning for conflicting statements detection in text](#)

[Lie-o-matic: using natural language processing to detect contradictory statements](#)

[Natural language inference | NLP-progress \(nlpprogress.com\)](#)

# Thank You

---

*Questions?*

## Meet the Team

Meet our team of four Halıcıoğlu Data Science Institute graduates who are now consulting with Deloitte. Our expertise ranges from machine learning to data visualization, and we are committed to staying at the forefront of machine learning research and are continuously expanding our knowledge by attending workshops, conferences, and collaborating with other data scientists.



**Emily Ramond**

eramond

Office: Seattle, WA  
Title: Analyst  
Offering: CBO



**Parker Addison**

paddison

Office: San Diego, CA  
Title: Consultant  
Offering: S&A



**Shweta Kumar**

shwetakumar5

Office: San Diego, CA  
Title: Analyst  
Offering: EP



**Max Levitt**

mlevitt

Office: San Diego, CA  
Title: Analyst  
Offering: CBO