

“It’s Legalese to Me:” Abstractive Summarization for Legislative Documents

Rebecca Nissan | rebeccanissan@berkeley.edu | W266: NLP | UC Berkeley School of Information

Abstract: Government legislation has a significant impact on people’s lives, but often lives in long, incomprehensible documents. This project aims to improve machine-generated summaries of legislation. With ROUGE scores as our primary metric, we explore (1) the impact of fine-tuning a generic transformer model (T5) for legal domain-specific summarization, and (2) the impact of pre-processing legislative documents prior to producing model-generated summaries. We find that our fine-tuned T5 model does not improve ROUGE scores relative to an out-of-the-box pre-trained T5 model. We also find that our pre-processing method leads to no improvement on ROUGE or summary quality. Finally, we discuss suggestions for future research, with an emphasis on task-specific pre-trained transformer models, other pre-processing methods, and alternative metrics for assessing semantic quality.

I. Introduction

Government legislation impacts our healthcare, education, infrastructure, and much more, but remains opaque to most people who lack a legal background and/or a lot of spare time. Worse yet, most people learn about legislation (and form opinions about it) through news outlets, which are often partisan.¹ Short, easily digestible summaries could help citizens learn about legislation without embedded partisan opinions or interpretations. This project aims to improve access to and understanding of legislation by improving the quality of machine-learning based summaries, which could ultimately serve as a step towards a less divisive political landscape.

II. Background

Many of the prior works in this domain begin by highlighting the challenging nature of the task, and for good reason: long documents and context-specific jargon make summarizing legislation particularly difficult, even for machines [1]. Many state-of-the-art summarization approaches are trained and tested on news articles, which contain less jargon and have a mean input text length that is “about 16 times shorter” than legislative bills [2]. Nonetheless, recent research demonstrates that transformer architectures can improve machine performance on legislative document summarization relative to unsupervised baselines [3]. The team that introduced BillSum (the dataset used for this project) established benchmarks for transformer performance on *extractive* summarization [4], but abstractive summarization has been shown to more closely match human summaries in a legal context [5]. Following the release of the BillSum dataset, others evaluated abstractive summarization methods to improve on the extractive baseline, showing better results. Perhaps the most promising of those experiments is PEGASUS, a transformer pre-trained for summarization tasks specifically, but it is trained on the C4 common crawl and HugeNews datasets, neither of which are context-specific for our task [6]. Fine-tuning a generic transformer-based language model for this task has also been explored – Gallegos and George (2022) found that fine-tuning improved the performance of Facebook’s BART-Large-CNN on the BillSum dataset. While Google’s T5 has been tested on this dataset without fine tuning [7], to the best of our knowledge this particular model has not been evaluated on this dataset with finetuning.

¹ Of course, this approach is not immune to partisanship if summaries used as labels are written through a partisan lens, but a qualitative review of our dataset strongly suggests that the target summaries in our dataset are neutral i.e. do not reflect the author’s opinions about the bill

Even with abstractive approaches, transformer architectures, and task-specific finetuning, summaries of such complex documents are likely to benefit from strategic pre-processing. Given that T5 and similar transformer models truncate input, and legislative documents are very long, it may be useful to create inputs that are themselves a kind of preliminary “summary,” capturing the main points from the whole document, not just the first subset. Klaus et al., who fine-tuned a pre-trained BERT model for summarization of European legislation, concluded that “a multistep approach is promising,” [3] but a recent attempt at document simplification prior to summary generation on BillSum found that adding this step did not improve results [1]. However, the authors note that the simplification algorithm used in this case (“[ACCESS](#)” by Facebook) has shortcomings: it is not trained on specific legal language, and it has length constraints. For these reasons, they only used the 10% of the BillSum data with the shortest input document lengths for the simplification and summarization experiment. This suggests it could be useful to explore other pre-processing or simplification steps prior to text generation in this domain, and that there is room for improvement on long input documents in particular.

II. Project Overview

Dataset

This project leverages the [BillSum](#) dataset, which was introduced by Kornilova and Eidelman (2019) [4]. BillSum is a corpus of legislative bills and target, human-written summaries. US federal bills are typically released along with human summaries, but state and local bills are not. This dataset leverages the target summaries released by the US government to train models that summarize legislative text, which can ultimately be used to generate summaries for state and local bills. The dataset includes a training set alongside two test sets: one with US bills and another with bills from the California State Legislature. The authors demonstrate that their models translate well from the US bills to the state ones. For our experiment, we split the training set into “train” and “validation” sets: 15,159 and 3,790 rows respectively. While ultimately the goal would be to test new methods on the state legislative data, as a first step we evaluate our models on a sample from of the 3,269 row US bills test set. For all subsets, each row contains the original text, the human-generated target summary, and the bill’s title. Our most notable insight from exploratory data analysis (and a review of the original BillSum paper) is that both the bills themselves and the target summaries are long: the dataset architects include only bills that are between 5,000 and 20,000 characters (which corresponds to an average of 1,285 words) and note that most target summaries of US bills range from 1,000 to 2,000 characters.

Evaluation Metrics

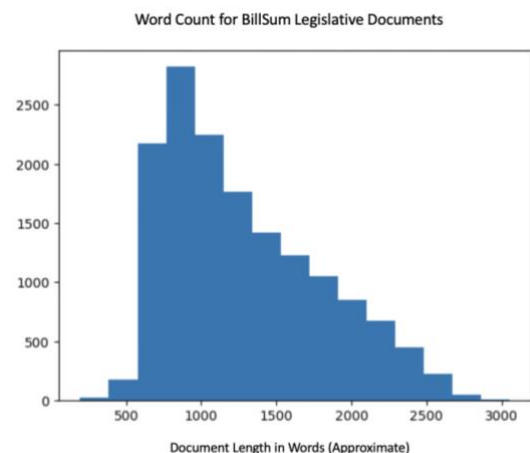


Figure 1: Histogram of word count for BillSum documents in training data (text inputs to our model).

Our primary evaluation metric for this task is ROUGE, which is the standard for evaluating model performance on summarization tasks. Fundamentally, ROUGE measures n-gram overlap between the target and candidate summaries. We choose this metric because it is standard and provides a useful benchmark, but then discuss ROUGE’s limitations and propose alternatives. In order to put our ROUGE performance in context, it is worth noting benchmarks from relevant prior work: the original BillSum paper, which fine-tuned BERT for extractive summarization using this dataset, achieved ROUGE-L scores in the 32-34 point range [4]. PEGASUS without fine-tuning achieved ~45 ROUGE-L points for abstractive summarization using BillSum [6]. We will also use a qualitative assessment of the output summaries to assess model performance: since our main goal is to make legislation more accessible to the public, it is important that our summaries are not just faithful to the targets, but also fluid and readable.

Baseline

For an initial, unsupervised baseline, we use SumBasic [8]. This extractive approach prioritizes sentences for the summary that contain words that are used frequently in the input. SumBasic is a useful baseline because it runs with limited computational resources and can handle long inputs. As an additional baseline, we use the base version of Google’s [T5](#), a pre-trained transformer that can be used for a variety of natural language processing tasks, including but not limited to summarization [9]. For the baseline assessment of T5 on this task, we do no pre-processing of the text and we use the default hyperparameters for model generation.

Modeling

We conduct three main experiments. All are evaluated on a randomly selected subset of 500 bill/target pairs from the test set.

Experiment 1: First, we modify T5’s hyperparameters for model generation. We experimented with several sets of hyperparameters on a small sample, and found that increasing `num_beams` and `no_repeat_ngram_size` did not qualitatively improve outputs (although we recommend revisiting this in future work). Most importantly, we noticed that output summaries using the default configuration were very short and seemed to end abruptly, so we increased the minimum length for output summaries from the default T5-base configuration of 30 tokens to 50. Due to limitations in compute, we were not able to set the `max_new_tokens` value beyond 150.

Experiment 2: We then fine-tune T5 by training it on the BillSum dataset (15,159 rows) for three epochs. Sparse categorical cross-entropy loss decreases with each epoch and reaches 0.80 by the final epoch, but validation loss increases after the second epoch to 1.44. (This may suggest some overfitting – potential shortcomings of our fine-tuning approach are discussed below).

Experiment 3: Finally, we pre-process a copy of the test subset ($n=500$) by leveraging patterns in the structure of the documents. Since T5 truncates inputs to 512 tokens, we hypothesize that taking key portions of text from throughout the document (and using those to create a shorter input for the model) would better capture the main points. The legislative bills are split into sections, which are all labeled with “Section” or “SEC.” For pre-processing, we take the n tokens after each instance of “SEC” or “Section” and concatenate them into a final input. The pre-processed test data is passed into the same fine-tuned model from experiment 2.

III. Results

We find that adjusting the T5 hyperparameters for text generation improves ROUGE performance compared to both baselines, even when using the out-of-the-box T5 model. Especially since T5 is such a generic language model, we would expect to see improved results with the fine-tuned model, but we do not see a significant difference between fine-tuned and out-of-the-box T5 with the hyperparameters held constant. Finally, our pre-processing scheme also does not appear to meaningfully impact the ROUGE score in either direction. ROUGE scores are reported in Table 1.

IV. Discussion of Results

The ROUGE score results and a qualitative analysis of the summary outputs reveal several insights. Firstly, it is worth noting that our SumBasic baseline is worse than the same metric reported in prior work using the same data (e.g. ROUGE-L of 23.92 in [4]), and we see qualitatively poor outputs such as “Section 1” (because this “sentence” appears very often in our inputs). We hypothesize that this is because we set the summary length to 5% of the original, which is probably far too short.

	BillSum Test Set n=500			
	ROUGE-1	ROUGE-2	ROUGE-L	ROUGE-L-SUM
Baseline (SumBasic)	9.74	4.20	7.28	8.56
T5 Base, Default Hyperparameters	10.66	4.24	8.90	9.21
T5 Base, Adjusted Hyperparameters	24.31	7.86	15.56	17.04
T5 Base Fine Tuned on BillSum, Adjusted Hyperparameters	21.22	7.13	15.70	16.90
T5 Base Fine Tuned on BillSum, Adjusted Hyperparameters, Pre-processed Inputs	22.61	9.14	17.47	18.65

Table 1: ROUGE Performance on BillSum Test Data: All Experiments

Secondly, our results highlight the importance of the hyperparameters for text generation: simply modifying these resulted in a 6.66 point improvement in ROUGE-L over the default configuration. We could likely improve results even further with additional tuning: for example, we could do experiments with many beams (we only tried up to 3). A review of the summary outputs (see Appendix 1) shows that summaries seem to stop abruptly or devolve into nonsense tokens like “rerere” by the end of generation. This matches prior findings that with BillSum in particular, “some predictions matched the beginning portion of a reference, but cut off prematurely to produce an incomplete summary.” [1] Future modifications (compute permitting) could increase both minimum and maximum lengths for the output summaries, although this is only likely to help if model inputs are less truncated or are pre-

processed to include the vital information from a long bill (otherwise the summaries will still reflect only the early parts of the bill). This truncation problem is likely one of the reasons our fine-tuning did not improve T5 performance. Due to resource constraints, we truncated input to 168 tokens (less than the maximum of 512). Overfitting to the training data is also possible, since we see loss increase over time on the validation set, even when the training loss continues to decrease.

Given the truncation constraints of many pre-trained models, we hypothesized that pre-processing the data to gather important information from throughout the document would improve performance. This also appears not to be the case, although a qualitative review of our pre-processed inputs suggests that our approach might be too simplistic: choosing an arbitrary number of tokens after the section header (even though we tried a few different values for this n) leaves many abruptly-ending sentences in the *input*. This limitation is especially significant given that we are using a language model, which is designed to predict the likelihood of masked tokens based on surrounding tokens – if some of the surrounding tokens in an input sentence suddenly come from a different portion of the bill entirely, the model is unlikely to produce comprehensible outputs. An alternative pre-processing approach that leverages text simplification techniques could still be promising, although the experiment with ACCESS in prior work suggests that simplification will be challenging unless the model is context-specific and designed to handle long inputs.

Finally, evaluating results for abstractive summarization tasks is itself a challenge, given the shortcomings of the ROUGE metric, which does not account for the fact that there are many ways to say the same thing, and a summary is not necessarily lower quality if it uses different words to convey the same meaning as the target [10]. In fact, citizens reading legislative summaries might actually prefer different phrases from those chosen in the targets, which were written by experts who read legislative documents frequently.

V. Follow-Up Experiments

In summary, our initial experiments had several shortcomings: (1) T5 does not perform very well out of the box since it is a generic model, and our fine-tuning did not improve results, (2) our pre-processing scheme appears to be too basic and (3) the metric used may not even capture summary quality adequately. To address these, we conduct the following analyses:

(1) Using a randomly selected small sample from our test set ($n=10$), we generate summaries using a version of PEGASUS that was fine-tuned on BillSum (from [HuggingFace](#)) and find the below results, suggesting that PEGASUS, especially fine-tuned for this domain, is a better model for the task since it is summarization-specific out of the box. Our hyperparameters for generation also do less well than PEGASUS’s defaults, further suggesting that our hyperparameters for T5 generation could be improved.

(2) As an alternative pre-processing step, we attempt to simplify the bill text, turning “legalese” into plain English. Unfortunately, to the best of our knowledge there is no available labeled dataset for simplification of legal text, but a proprietary solution does exist: “[Legalese Decoder](#).” According to their documentation, the app’s algorithm first looks up legal concepts in a database and then “rephrases the document” using simpler language.

	BillSum Test Set n=10			
	<i>ROUGE-1</i>	<i>ROUGE-2</i>	<i>ROUGE-L</i>	<i>ROUGE-L-SUM</i>
PEGASUS (Adjusted hyperparameters)	59.21	37.72	46.36	48.67
PEGASUS (Default hyperparameters)	60.70	41.61	49.37	51.18

Table 2: ROUGE Performance on BillSum Test Data with PEGASUS

While we do not know exactly what steps are being taken to simplify inputs, we consider it a useful exercise to pass “legalese-decoded” bills into our algorithm and assess performance. This was done manually (they do not have an API), so we use only 5 examples. ROUGE scores do not improve, nor does summary readability (see Appendix 2).

	BillSum Test Set (n=5)			
	<i>ROUGE-1</i>	<i>ROUGE-2</i>	<i>ROUGE-L</i>	<i>ROUGE-L-SUM</i>
T5 Base Fine Tuned on BillSum, Adjusted Hyperparameters <i>Original Inputs</i>	27.17	6.40	17.23	18.71
T5 Base Fine Tuned on BillSum, Adjusted Hyperparameters <i>Legalese Decoded Inputs</i>	25.78	7.31	17.37	19.34

Table 3: ROUGE Performance on BillSum Test Data with Legalese-Decoded Inputs

(3) Finally, we conduct a small experiment to see whether an alternative measure of similarity between the target and model-generated summaries would more closely match human interpretations of summary quality. For 5 examples, we produce Sentence-BERT embeddings for target and output summaries, and then calculate the cosine similarity between those embeddings. Sentence-BERT is chosen because it uses sentence-level embeddings and, compared to ROUGE, may help us better account for different ways to convey the same meaning [11]. We then ask a human grader blind to the experiment to rank each output summary for both readability and perceived similarity to the target. This initial experiment (results included in Appendix 3) does not show any meaningful pattern, but we recommend conducting a similar test at larger scale to more fully assess this potential alternative way to measure summary quality.

VI. Conclusion & Directions for Future Research

In summary, context-specificity and document length remain significant challenges for summarizing legislation. We found that T5, even when fine-tuned on for BillSum summarization, generates poor ROUGE scores and qualitatively weak summaries. PEGASUS, which is pre-trained specifically for

summarization, performs far better at this task. Two preliminary attempts at pre-processing inputs prior to generating summaries were not especially promising, but additional work should continue to explore this question at greater scale and with more sophisticated methods – for example, a very strong extractive summary could be fed into an abstractive summarization model. Specifically, it would be interesting to simplify or pre-process inputs to the fine-tuned PEGASUS model and evaluate whether this improves output readability. In order to effectively test simplification processes, it would be valuable to create an open-source labeled dataset with simplified long-form versions of legislative texts. Other approaches for dealing with long and complex text (such as pointer-generators) could also be borrowed from non-legal domains [12]. Finally, without a better metric for assessing the quality of abstractive summaries, it will be difficult to meaningfully advance research in this area and improve output for end users (human beings who want fluid text). To this end, we suggest follow-up research comparing different possible metrics with human evaluations and determining which metrics (or combination of metrics) best capture summary quality.

REFERENCES

- [1] Gallegos, I. & George, K. (2022) The Right to Remain Plain: Summarization and Simplification of Legal Documents. *Stanford CS224N Custom Project*.
https://web.stanford.edu/class/cs224n/reports/custom_116652906.pdf
- [2] Aumiller, D., Couhan, A., and Gertz, M. (2022) EUR-Lex-Sum: A Multi- and Cross-lingual Dataset for Long-form Summarization in the Legal Domain. *CoRR*. <https://arxiv.org/pdf/2210.13448.pdf>
- [3] Klaus, S. et al. (2022) Summarizing Legal Regulatory Documents Using Transformers. *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*.
<https://api.semanticscholar.org/CorpusID:250340205>
- [4] Kornilova, A., and Eidelman, V. (2019). BillSum: A Corpus for Automatic Summarization of US Legislation. *Proceedings of the 2nd Workshop on New Frontiers in Summarization*, pages 48–56, Hong Kong, China. Association for Computational Linguistics. <https://aclanthology.org/D19-5406.pdf>
- [5] Feijo, D.V., & Moriera, V.P. (2019) Summarizing Legal Rulings: Comparative Experiments. *Proceedings of Recent Advances in Natural Language Processing*, pages 313-322, Varna, Bulgaria, Sep 2-4, 2019.
<https://acl-bg.org/proceedings/2019/RANLP%202019/pdf/RANLP036.pdf>
- [6] Zhang, J., Zhao, Y., Saleh, M., and Liu, P.J. (2020) PEGASUS: Pre-Training with Extracted Gap-sentences for Abstractive Summarization. *Proceedings of the 37th International Conference on Machine Learning*, Vienna, Austria, PMLR 119, 2020. <https://arxiv.org/pdf/1912.08777.pdf>

- [7] Rehman, T., Das, S., Sanyal, D.K., and Chattopadhyay, S. (2023) An Analysis of Abstractive Text Summarization Using Pre-trained Models. *Proceedings of International Conference on Computational Intelligence, Data Science, and Cloud Computing*, pages 253-264. <https://arxiv.org/pdf/2303.12796.pdf>
- [8] Nenkova, A. & Vanderwende, L. (2005) The Impact of Frequency on Summarization. <https://www.cs.bgu.ac.il/~elhadad/nlp09/sumbasic.pdf>
- [9] Raffel, C., et al. (2020) Exploring the Limits of Transfer Learning with a Unified Text-to-Text Transformer. *Journal of Machine Learning Research*, 21, 1-67. <https://arxiv.org/pdf/1910.10683.pdf>
- [10] Shluter, N. (2007) The Limits of Automatic Summarisation According to ROUGE. *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics*, Volume 2, pages 41-45. <https://aclanthology.org/E17-2007/>
- [11] Reimers, N., & Gurevych, I. (2019) Sentence-BERT: Sentence Embeddings using Siamese BERT-Networks. *EMNLP 2019*. <https://arxiv.org/abs/1908.10084>
- [12] See, A., Liu, P., and Manning, C.D. (2017) Get to the Point: Summarization with Pointer-Generator Networks. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, Volume 1, pages 1073-1083. <https://aclanthology.org/P17-1099/>

APPENDIX 1

An example of summary outputs for a single bill across baselines and experiments. Illustrates the shortcomings of some of our models. Note especially “T5 Base, Adjusted Hyperparameters,” where the model hallucinates and inserts details about “John McCain” and “Obama,” who are not mentioned in the original text. Note also that the T5 fine-tuned summaries start off on the right track, but then become nonsensical and end abruptly. Finally, note that PEGASUS produces much cleaner summaries, but they still seem to end abruptly.

Title	Forest Service Financial Accountability Restoration Act of 1999
Original Text	<p>SECTION 1. SHORT TITLE.\n\n This Act may be cited as the ``Forest Service Financial Accountability Restoration Act of 1999".\n\nSEC. 2. USE OF INCENTIVES TO RESTORE FOREST SERVICE FINANCIAL ACCOUNTABILITY.\n\n (a) Findings.--The Congress finds the following:\n (1) Section 3521(e) of title 31, United States Code, \n requires the Inspector General of the Department of Agriculture \n to audit the annual financial statements required to be \n prepared by the Forest Service. As a result of the inability of \n the Forest Service to produce reliable financial statements, \n the agency has received four successive adverse opinions or \n disclaimers from the Inspector General.\n (2) In January 1999, the Comptroller General submitted a \n report to Congress that identified the Forest Service as an \n agency at ``high risk" of waste, fraud, abuse, and \n mismanagement due to the persistent and severe weaknesses in \n the Forest Service's accounting and financial reporting \n systems.\n (3) The Comptroller General has stated that the Forest \n Service will maintain its high risk status until the agency, at \n a minimum--\n (A) receives two consecutive unqualified audit \n opinions from the Inspector General; and\n (B) corrects the material internal control \n weaknesses, identified in the Inspector General's audit \n reports of the financial statements of the Forest \n Service, which adversely affect the ability of the \n Forest Service to maintain accountability over its \n assets on an ongoing basis.\n (4) Despite initial efforts on the part of the Forest \n Service to correct deficiencies in its financial management \n systems, the Comptroller General and the Inspector General of \n the Department of Agriculture have indicated that the Forest \n Service may still be years away from producing reliable \n financial statements or changing its high risk status.\n (5) Performance incentives are necessary to ultimately \n restore financial accountability to the Forest Service, and \n such incentives must be structured so as to preserve the \n ability of the Forest Service to perform its core missions, \n particularly a sustained reduction of the natural resources \n restoration and maintenance backlog within the National Forest \n System, while safeguarding further investments of taxpayer \n dollars from waste, fraud, abuse, and mismanagement.\n (b) Conditional Limitations on Forest Service Appropriations.--\nSubject to subsections (c) and (d), for fiscal years beginning after \n the date of the enactment of this Act, the total amount appropriated \n for the Forest Service to carry out discretionary programs and \n activities for a fiscal year shall not exceed the sum of the amounts \n appropriated under the heading ``Forest</p>

	<p>Service" in title II of the \nDepartment of the Interior and Related Agencies Appropriations Act, \n1999 (as contained in section 101(e) of division A of Public Law 105-\n277; 112 Stat. 2681-268).\n (c) Removal of Condition.--The limitations on the level of Forest \nService appropriations or obligations specified in subsection (b) shall \nterminate on the earlier of the following:\n (1) The date on which the Forest Service is no longer an \n agency at high risk of waste, fraud, abuse, and mismanagement \n due to weaknesses in its accounting and financial reporting \n systems, as determined under subsection (d).\n (2) The end of the fifth fiscal year subject to the \n limitations.\n (d) Criteria for Removal of High Risk Status.--The Forest Service \nshall no longer be an agency at high risk under subsection (c)(1) if \nthe Comptroller General certifies in writing to Congress that Forest \nService financial management is no longer at high risk. The Comptroller \nGeneral may make that certification only if--\n (1) the Forest Service receives two consecutive unqualified \n audit opinions from the Inspector General of the Department of \n Agriculture under section 3521(e) of title 31, United States \n Code; and\n (2) the Forest Service corrects the material weaknesses \n identified in the audit reports of the Inspector General \n regarding the Forest Service's fiscal year 1995 financial \n statements and all subsequent audit reports of the Inspector \n General regarding the Forest Service's financial statements \n published before the date of the enactment of this Act.\n (e) Backlog Reduction Outputs.--During each fiscal year identified \nin subsection (c)(2), the Forest Service shall maintain or increase \noutputs, relative to the levels achieved during fiscal year 1999, from \nprograms that directly reduce the natural resources restoration and \nmaintenance backlog within the National Forest System. Such outputs \nshall include the following:\n (1) Acres of forested land treated for reduction of \n wildfire risk.\n (2) Acres of forested land treated for reduction of \n mortality risk from insect and disease infestation.\n (3) Miles of roads and trails reconstructed or maintained.\n (4) Number of recreational facilities reconstructed or \n maintained.\n (5) Acres of terrestrial habitat restored or enhanced.\n (6) Miles of riparian areas restored or enhanced.\n (7) Other priority outputs identified by the Secretary.</p>
Target	<p>Forest Service Financial Accountability Restoration Act of 1999 - Places specified limits on Forest Service appropriations until the earlier of the following: (1) the Service is no longer an agency at high risk of waste, fraud, and mismanagement due to accounting and financial reporting weaknesses; or (2) the end of the fifth fiscal year subject to such limitations. \nEnumerates the criteria under which the Comptroller General shall certify to Congress that the Service is no longer at high risk \nSets forth Service backlog reduction requirements.</p>
SumBasic Baseline	<p>2) Acres of forested land treated for reduction of \n mortality risk from insect and disease infestation.</p>
T5 Base, Default HPs	<p>a report to congress identified the Forest Service as at 'high risk' of waste,</p>
T5 Base, Adjusted HPs	<p>a report to congress identified the Forest Service as at 'high risk' of waste, fraud, abuse. the Comptroller general has stated that the agency will maintain its high risk status if it corrects deficiencies - despite initial efforts... obama has not acted on this</p>

	issue he says ; sen. john mccain says the act is ok unless it is repealed, but not before january 1, 1999 eve dre
T5 Base Fine Tuned on BillSum, Adjusted HPs	Amends the Forest Service Organic Act to prohibit the Secretary of Agriculture from acquiring, supplying, or delivering for use in any Forest service program rere a forest service that has incurred the cost of establishing or maintaining
T5 Base Fine Tuned on BillSum, Adjusted HPs, Pre-processed Inputs	Requires the Secretary of Agriculture to establish a Forest Service financial assistance program to provide financial support to forest service organizations and other eligible entities that serve rere s appropriations. Authorizes an apportion to
PEGASUS (Adjusted hyperparameters)	Forest Service Financial Accountability Restoration Act of 1999 - States that specified limitations on Forest service appropriations or obligations shall terminate on the earlier of: (1) the date on which the Forest Services is no longer an agency at high risk of waste, fraud, abuse, and mismanagement due to weaknesses in its accounting and financial reporting systems; and (2) the end of the fifth fiscal year subject to such limitations. <n>Declares that the FOREST Service shall no more be an entity athigh risk if the Comptroller General certifies in writing to the Congress that Forest SERVICE financial management is No longer at such high Risk. Sets forth conditions under which The Comptroller general may certify such a removal. (Sec. 3) Requires the forest service to maintain or increase outputs, relative to those achieved during
PEGASUS (Default hyperparameters)	Forest Service Financial Accountability Restoration Act of 1999 - States that specified limitations on Forest Service appropriations or obligations shall terminate on the earlier of: (1) the date on which the Forest Service is no longer an agency at high risk of waste, fraud, abuse, and mismanagement due to weaknesses in its accounting and financial reporting systems; and (2) the end of the fifth fiscal year subject to such limitations. <n>Declares that the Forest Service shall no longer be at high risk if the Comptroller General certifies in writing to the Congress that Forest Service financial management is no longer at high risk. Authorizes the Comptroller General to make such certification only if: (1) the Forest Service receives two consecutive unqualified audit opinions from the Inspector General of the Department of Agriculture; and (2) the Forest Service corrects the material weaknesses identified in the audit reports of the Inspector General regarding the Forest Service's FY 1995 financial statements and all subsequent audit reports of the Inspector General regarding such statements published before the date of the enactment of this Act. <n>Requires the Forest Service, during each fiscal year, to maintain or increase output, relative to the levels achieved during FY 1999, from programs that directly reduce the natural resources restoration and maintenance backlog within the National Forest System, including: (1) acres of forested land treated for reduction of wildfire risk

APPENDIX 2

An example of simplified (legalese-decoded) outputs. Note that the simplified input is much easier to read, but still long, and its summary is not meaningfully more readable or closer to the target compared with the summary generated from the original input. This is likely because the model is generally not performing well, so we are unlikely to see significant gains by changing the input text.

Title	Fair Access to Investment Research Act of 2016
Original Text	<p>SECTION 1. SHORT TITLE. This Act may be cited as the "Fair Access to Investment Research Act of 2016".</p> <p>SEC. 2. SAFE HARBOR FOR INVESTMENT FUND RESEARCH.</p> <p>(a) Expansion of the Safe Harbor.--Not later than the end of the 45-day period beginning on the date of enactment of this Act, the Securities and Exchange Commission shall propose, and not later than the end of the 180-day period beginning on such date, the Commission shall adopt, upon such terms, conditions, or requirements as the Commission may determine necessary or appropriate in the public interest, for the protection of investors, and for the promotion of capital formation, revisions to section 230.139 of title 17, Code of Federal Regulations, to provide that a covered investment fund research report that is published or distributed by a broker or dealer-- (1) shall be deemed, for purposes of sections 2(a)(10) and 5(c) of the Securities Act of 1933 (15 U.S.C. 77b(a)(10), 77e(c)), not to constitute an offer for sale or an offer to sell a security that is the subject of an offering pursuant to a registration statement that is effective, even if the broker or dealer is participating or will participate in the registered offering of the covered investment fund's securities; and (2) shall be deemed to satisfy the conditions of paragraph (a)(1) or (a)(2) of section 230.139 of title 17, Code of Federal Regulations, or any successor provisions, for purposes of the Commission's rules and regulations under the Federal securities laws and the rules of any self-regulatory organization.</p> <p>(b) Implementation of Safe Harbor.--In implementing the safe harbor pursuant to subsection (a), the Commission shall-- (1) not, in the case of a covered investment fund with a class of securities in substantially continuous distribution, condition the safe harbor on whether the broker's or dealer's publication or distribution of a covered investment fund research report constitutes such broker's or dealer's initiation or reinitiation of research coverage on such covered investment fund or its securities; (2) not-- (A) require the covered investment fund to have been registered as an investment company under the Investment Company Act of 1940 (15 U.S.C. 80a-1 et seq.) or subject to the reporting requirements of section 13 or 15(d) of the Securities Exchange Act of 1934 (15 U.S.C. 78m, 78o(d)) for any period exceeding the period of time referenced under paragraph (a)(1)(i)(A)(1) of section 230.139 of title 17, Code of Federal Regulations; or (B) impose a minimum float provision exceeding that referenced in paragraph (a)(1)(i)(A)(1)(i) of section 230.139 of title 17, Code of Federal Regulations; (3) provide that a self-regulatory organization may not maintain or enforce any rule that would-- (A) prohibit the ability of a member to publish or distribute a covered investment fund research report solely because the member is also participating in a registered offering or other distribution of any securities of such covered investment fund; or (B) prohibit the ability of a member to participate</p>

	<p>in a registered offering or other distribution of securities of a covered investment fund solely because the member has published or distributed a covered investment fund research report about such covered investment fund or its securities; and (4) provide that a covered investment fund research report shall not be subject to section 24(b) of the Investment Company Act of 1940 (15 U.S.C. 80a-24(b)) or the rules and regulations thereunder, except that such report may still be subject to such section and the rules and regulations thereunder to the extent that it is otherwise not subject to the content standards in the rules of any self-regulatory organization related to research reports, including those contained in the rules governing communications with the public regarding investment companies or substantially similar standards. (c) Rules of Construction.--Nothing in this Act shall be construed as in any way limiting-- (1) the applicability of the antifraud or antimanipulation provisions of the Federal securities laws and rules adopted thereunder to a covered investment fund research report, including section 17 of the Securities Act of 1933 (15 U.S.C. 77q), section 34(b) of the Investment Company Act of 1940 (15 U.S.C. 80a-33(b)), and sections 9 and 10 of the Securities Exchange Act of 1934 (15 U.S.C. 78i, 78j); or (2) the authority of any self-regulatory organization to examine or supervise a member's practices in connection with such member's publication or distribution of a covered investment fund research report for compliance with applicable provisions of the Federal securities laws or self-regulatory organization rules related to research reports, including those contained in rules governing communications with the public. (d) Interim Effectiveness of Safe Harbor.-- (1) In general.--From and after the 180-day period beginning on the date of enactment of this Act, if the Commission has not adopted revisions to section 230.139 of title 17, Code of Federal Regulations, as required by subsection (a), and until such time as the Commission has done so, a broker or dealer distributing or publishing a covered investment fund research report after such date shall be able to rely on the provisions of section 230.139 of title 17, Code of Federal Regulations, and the broker or dealer's publication of such report shall be deemed to satisfy the conditions of paragraph (a)(1) or (a)(2) of that section if the covered investment fund that is the subject of such report satisfies the reporting history requirements (without regard to Form S-3 or Form F-3 eligibility) and minimum float provisions of such subsections for purposes of the Commission's rules and regulations under the Federal securities laws and the rules of any self-regulatory organization, as if revised and implemented in accordance with subsections (a) and (b). (2) Status of covered investment fund.--After such period and until the Commission has adopted revisions to section 230.139 of title 17, Code of Federal Regulations, and FINRA has revised rule 2210, for purposes of subsection (c)(7)(O) of such rule, a covered investment fund shall be deemed to be a security that is listed on a national securities exchange and that is not subject to section 24(b) of the Investment Company Act of 1940 (15 U.S.C. 80a-24(b)). Communications concerning only covered investment funds that fall within the scope of such section shall not be required to be filed with FINRA. (e) Definitions.--For purposes of this Act: (1) The term "covered investment fund research report" means a research report published or distributed by a broker or dealer about a covered investment fund or any securities issued by the covered investment fund, but does not include a research report to the extent that the research report is published or distributed by the covered investment fund or any affiliate of the covered investment fund. (2) The term "covered investment fund"</p>
--	--

	<p>means-- (A) an investment company registered under, or that has filed an election to be treated as a business development company under, the Investment Company Act of 1940 (15 U.S.C. 80a-1 et seq.) and that has filed a registration statement under the Securities Act of 1933 (15 U.S.C. 77a et seq.) for the public offering of a class of its securities, which registration statement has been declared effective by the Commission; and (B) a trust or other person-- (i) issuing securities in an offering registered under the Securities Act of 1933 and which class of securities is listed for trading on a national securities exchange; (ii) the assets of which consist primarily of commodities, currencies, or derivative instruments that reference commodities or currencies, or interests in the foregoing; and (iii) that provides in its registration statement under the Securities Act of 1933 that a class of its securities are purchased or redeemed, subject to conditions or limitations, for a ratable share of its assets. (3) The term ``FINRA'' means the Financial Industry Regulatory Authority. (4) The term ``research report'' has the meaning given that term under section 2(a)(3) of the Securities Act of 1933 (15 U.S.C. 77b(a)(3)), except that such term shall not include an oral communication. (5) The term ``self-regulatory organization'' has the meaning given to that term under section 3(a)(26) of the Securities Exchange Act of 1934 (15 U.S.C. 78c(a)(26)).</p>
Simplified Text (Legalese-Decoded)	<p>SECTION 1. SHORT TITLE This law is called the "Fair Access to Investment Research Act of 2016".</p> <p>SEC. 2. SAFE HARBOR FOR INVESTMENT FUND RESEARCH</p> <p>(a) Expansion of the Safe Harbor The government agency in charge of regulating securities and investments (called the Securities and Exchange Commission) has to make some changes to a specific regulation within 180 days of this law being passed. These changes are meant to protect investors and promote the creation of new investments. The new regulation will say that a report about a specific investment fund, written by a broker or dealer, will not be considered a sale offer or a solicitation for that investment fund's securities, even if the broker is involved in selling those securities. It will also say that the report meets certain conditions set by the Commission's rules and regulations.</p> <p>(b) Implementation of Safe Harbor The government agency will also have to figure out how to enforce this new safe harbor regulation.</p> <p>According to this rule (subsection (a)), the Commission is not allowed to make certain conditions for a safe harbor related to a type of investment fund that is continuously distributing securities. They cannot make it a requirement for a broker or dealer to publish or distribute a research report on this investment fund in order to initiate or restart research coverage on the fund or its securities. The Commission also cannot require the investment fund to be registered as an investment company or subject to certain reporting requirements for a longer period of time than what is mentioned in another regulation. They also cannot impose a minimum provision that exceeds what is mentioned in that regulation. Furthermore, a self-regulatory organization cannot have a rule that prevents a member from publishing or distributing a research report on the investment fund solely because the member is participating in a registered offering or distribution of any securities of that investment fund. The organization also cannot prevent a member from participating in a registered offering or distribution of securities of the investment fund solely because the member has published or distributed a research report on that investment fund or its securities. Lastly, a research report on the investment fund is not subject to a specific section of another law and its regulations, except if it does</p>

	<p>not meet the content standards set by self-regulatory organizations regarding research reports on investment companies or similar standards.</p> <p>(c) Rules of Construction.--Nothing in this Act shall be construed as in any way limiting--</p> <p>(1) the applicability of the antifraud or antimanipulation provisions of the Federal securities laws and rules adopted thereunder to a covered investment fund research report, including section 17 of the Securities Act of 1933 (15 U.S.C. 77q), section 34(b) of the Investment Company Act of 1940 (15 U.S.C. 80a-33(b)), and sections 9 and 10 of the Securities Exchange Act of 1934 (15 U.S.C. 78i, 78j); or</p> <p>(2) the authority of any self-regulatory organization to examine or supervise a member's practices in connection with such member's publication or distribution of a covered investment fund research report for compliance with applicable provisions of the Federal securities laws or self-regulatory organization rules related to research reports, including those contained in rules governing communications with the public.</p> <p>(d) Interim Effectiveness of Safe Harbor.--</p> <p>(1) In general.-- From and after the 180-day period beginning on the date of enactment of this Act, if the Commission has not adopted revisions to section 230.139 of title 17, Code of Federal Regulations, as required by subsection (a), and until such time as the Commission has done so, a broker or dealer distributing or publishing a covered investment fund research report after such date shall be able to rely on the provisions of section 230.139 of title 17, Code of Federal Regulations, and the broker or dealer's publication of such report shall be deemed to satisfy the conditions of--</p> <p>Paragraph (a)(1) or (a)(2) of that section if the covered investment fund that is the subject of such report satisfies the reporting history requirements (without regard to Form S-3 or Form F-3 eligibility) and minimum float provisions of such subsections for purposes of the Commission's rules and regulations under the Federal securities laws and the rules of any self-regulatory organization, as if revised and implemented in accordance with subsections (a) and (b).</p> <p>Translation: If the investment fund being talked about in the report meets certain requirements set by the government and other organizations, then paragraph (a)(1) or (a)(2) of that section applies.</p> <p>(2) Status of covered investment fund.--After such period and until the Commission has adopted revisions to section 230.139 of title 17, Code of Federal Regulations, and FINRA has revised rule 2210, for purposes of subsection (c)(7)(O) of such rule, a covered investment fund shall be deemed to be a security that is listed on a national securities exchange and that is not subject to section 24(b) of the Investment Company Act of 1940 (15 U.S.C. 80a-24(b)). Communications concerning only covered investment funds that fall within the scope of such section shall not be required to be filed with FINRA.</p> <p>Translation: Until certain changes are made by the government and other organizations, a covered investment fund is considered to be a type of investment that is listed on a national exchange and not subject to certain rules. People don't have to submit certain documents to FINRA if they are only talking about these specific investment funds.</p> <p>(e) Definitions.--For purposes of this Act:</p> <p>(1) The term "covered investment fund research report" means a research report published or distributed by a broker or dealer about a covered investment fund or any securities issued by the covered investment fund, but does not include a research report to the extent that the research report is published or distributed by the covered investment fund or any affiliate of the covered investment fund.</p> <p>Translation: In this law, the term "covered investment fund research report" refers to a report made by a broker or dealer about a certain type of investment fund or the securities it issues. However, it</p>
--	--

	<p>doesn't include reports made by the investment fund itself or any of its related companies.</p> <p>(2) The term "covered investment fund" means: (A) a company that invests money for people and has registered with the government or chosen to be treated as a special type of company, and has also filed paperwork to sell a type of investment to the public. The government has approved this paperwork; (B) a trust or other organization: (i) that sells investments to the public and has registered with the government, and the investments can be bought and sold on a special market; (ii) whose main assets are things like goods, money from different countries, or special contracts based on goods or money; (iii) that says in its paperwork filed with the government that the investments it sells can be bought or sold, but only under certain conditions.</p> <p>(3) The term "FINRA" means the people who make sure that companies in the financial industry follow the rules.</p> <p>(4) The term "research report" means a document that gives information about an investment, as defined in a law, but does not include spoken information.</p> <p>(5) The term "self-regulatory organization" means a group that makes sure that companies in the investment market follow the rules, as defined in a law.</p>
Target	<p>Fair Access to Investment Research Act of 2016 This bill directs the Securities and Exchange Commission (SEC) to revise a specified regulation to create a safe harbor for certain publications or distributions of research reports by brokers or dealers distributing securities. The revised regulation shall declare that a published or distributed investment fund research report shall not be deemed to constitute an offer for sale or an offer to sell a security that is the subject of the offering pursuant to an effective registration statement, even if the broker or dealer is participating or will participate in the registered offering of the covered investment fund's securities. The covered investment fund research report shall indeed be deemed to satisfy the regulation's requirements as well as those of any self-regulatory organization. The SEC shall not impose specified conditions and requirements when implementing the safe harbor. Until the SEC has adopted such safe harbor revisions, and the Financial Industry Regulatory Authority (FINRA) has revised a related rule, a covered investment fund shall be deemed to be a security listed on a national securities exchange that is not subject to a certain filing requirement of the Investment Company Act of 1940.</p>
Original Summary (using T5 Base Fine Tuned on BillSum, Adjusted HPs)	<p>Amends the Internal Revenue Code to: (1) extend through December 31, 2015, the tax exemption for covered investment in mutually insured investment trusts; (2) reexports insured insurance companies; and (2) make permanent the investment tax credit</p>
Summary of Legalese- Decoded Input	<p>Amends the Internal Revenue Code to: (1) extend through December 31, 2017 the tax deduction for investment in a covered investment fund; (2) make permanent the reexempt investment account; and (3) make the minimum distribution of such fund available</p>

APPENDIX 3

ROUGE scores, sentence-BERT cosine similarity scores, and human rankings for 5 examples of target/candidate pairs. Use index to see target/candidate pairs for each row in the ranking / metrics table. No clear patterns are evident from this small sample, but this could provide directions for future work aiming to improve metrics for summarization.

Index	Title	Human Rankings (1 = low, 5 = high)	ROUGE-L	Cosine Sim. (SBERT)
1	To make technical corrections to the Water Resources Development Act of 1999.	Similarity: 2 / Readability: 1	9.20	57.30
2	Federal Forage Fee Act of 1993	Similarity: 3 / Readability: 4	24.69	68.35
3	Merchant Marine of World War II Congressional Gold Medal Act	Similarity: 2 / Readability: 2	19.20	43.59
4	To amend the Internal Revenue Code of 1986 to provide for unified income taxation with respect to pass-thru entities.	Similarity: 1 / Readability: 2	24.59	48.60
5	Fair Access to Investment Research Act of 2016	Similarity: 2 / Readability: 3	8.47	40.71

Index	Target/Candidates
1	<p>Target: Amends the Water Resources Development Act of 1999 to: (1) authorize appropriations for FY 1999 through 2009 for implementation of a long-term resource monitoring program with respect to the Upper Mississippi River Environmental Management Program (currently, such funding is designated for a program for the planning, construction, and evaluation of measures for fish and wildlife habitat rehabilitation and enhancement); (2) authorize the Secretary of the Army to carry out modifications to the navigation project for the Delaware River, Pennsylvania and Delaware, if such project as modified is technically sound, environmentally (currently, economically) acceptable, and economically justified; (3) subject certain previously deauthorized water resources development projects to the seven-year limitation governing project deauthorizations under the Act, with the exception of such a project for Indian River County, Florida; (4) except from a certain schedule of the non-Federal cost of the periodic nourishment of shore protection projects constructed after December 31, 1999, those projects for which a District Engineer's Report has been completed by such date; (5) require that the project cooperation agreement for the Comite River Diversion Project for flood control include a provision that specifies that any reduction in the non-Federal share</p>

	<p>that results from certain modifications be credited toward the share of project costs to be paid by the Amite River Basin Drainage and Water Conservation District; (6) allow the Secretary to provide additional compensation to Chesapeake City, Maryland (currently, to the City of Chesapeake, Maryland) for damage to its water supply resulting from the Chesapeake and Delaware Canal Project; (7) provide for the submission of certain reports on water resources development projects by the Secretary, notwithstanding Federal reporting termination provisions; and (8) authorize and provide for an authorization of appropriations for the existing program for the safety and operations expenses of the Federal Railroad Administration, and make available for obligation funds currently appropriated for such program.</p> <p>Candidate: Amends the Federal Water Pollution Control Act (commonly known as the Clean Water Act) to authorize the use of a water pollution control reresponsible to the holder of the permit for wetlands to be used</p>
2	<p>Target: Federal Forage Fee Act of 1993 - Subjects grazing operations on Federal land to applicable Federal, State, and local environmental and land use requirements. Sets forth a forage fee formula for lands under the jurisdiction of the Department of Agriculture and the Department of the Interior.</p> <p>Candidate: Amends the Federal Land Policy and Management Act of 1976 to require the Secretary of Agriculture to determine the cost of grazing on public lands, rere a management plan for the Area's acreage, and other</p>
3	<p>Target: . Merchant Marine of World War II Congressional Gold Medal Act (Sec. 3) This bill requires the Speaker of the House of Representatives and the President pro tempore of the Senate to arrange for the award, on behalf of Congress, of a single gold medal to the U.S. Merchant Marine of World War II, in recognition of their dedicated and vital service during World War II. Following its award the medal shall be given to the American Merchant Marine Museum where it will be available for display and research.</p> <p>Candidate: Directs the Secretary of Defense to establish a National Merchant Marine Corps of military retired pay to the Marines of the Korean War. Direct the Maritimes Commission to to determine the cost of such restitution. Requires</p>
4	<p>Target: Small Business Modernization Act of 2004 - Amends the Internal Revenue Code to: (1) terminate subchapter S corporation elections after 2004 and subchapter S status after 2014 and to allow privately-held domestic corporations, in lieu of electing subchapter S treatment, to elect to be treated as partnerships for tax purposes; (2) set forth rules for the tax treatment of former subchapter S corporations electing partnership status; and (3) exclude from net earnings from self-employment partnership income attributable to capital.</p> <p>Candidate: Amends the Internal Revenue Code to: (1) allow a tax deduction for the cost of certain transportation and related transportation, (2) make permanent the tax credit, and (3) the distribution of profits to taxable other than transportation to and from other sources</p>
5	<p>Target: Fair Access to Investment Research Act of 2016 This bill directs the Securities and Exchange Commission (SEC) to revise a specified regulation to create a safe harbor for certain publications or distributions of research reports by brokers or dealers distributing securities. The revised regulation shall declare that a published or distributed investment fund research report shall not be deemed to constitute an offer for sale or an offer to sell a security that is</p>

	<p>the subject of the offering pursuant to an effective registration statement, even if the broker or dealer is participating or will participate in the registered offering of the covered investment fund's securities. The covered investment fund research report shall indeed be deemed to satisfy the regulation's requirements as well as those of any self-regulatory organization. The SEC shall not impose specified conditions and requirements when implementing the safe harbor. Until the SEC has adopted such safe harbor revisions, and the Financial Industry Regulatory Authority (FINRA) has revised a related rule, a covered investment fund shall be deemed to be a security listed on a national securities exchange that is not subject to a certain filing requirement of the Investment Company Act of 1940.</p> <p>Candidate: Amends the Internal Revenue Code to: (1) extend through December 31, 2015, the tax exemption for covered investment in mutually insured investment trusts; (2) reexports insured insurance companies; and (2) make permanent the investment tax credit</p>
--	--