

War of Words II: Enriched Models of Law-Making Processes

Victor Kristof*

Ecole polytechnique fédérale de Lausanne

Matthias Grossglauser

Ecole polytechnique fédérale de Lausanne

Aswin Suresh*

Ecole polytechnique fédérale de Lausanne

Patrick Thiran

Ecole polytechnique fédérale de Lausanne

ABSTRACT

The European Union law-making process is an instance of a peer-production system. We mine a rich dataset of law edits and introduce models predicting their adoption by parliamentary committees. Edits are proposed by parliamentarians, and they can be in conflict with edits of other parliamentarians and with the original proposition in the law. Our models combine three different categories of features: (a) *Explicit* features extracted from data related to the edits, the parliamentarians, and the laws, (b) *latent* features that capture bi-linear interactions between parliamentarians and laws, and (c) *text* features of the edits. We show experimentally that this combination enables us to accurately predict the success of the edits. Furthermore, it leads to model parameters that are interpretable, hence provides valuable insight into the law-making process.

KEYWORDS

discrete choice models, natural language processing, computational social science, data-driven political science, web mining

ACM Reference Format:

Victor Kristof, Aswin Suresh, Matthias Grossglauser, and Patrick Thiran. 2021. War of Words II: Enriched Models of Law-Making Processes. In *Proceedings of the Web Conference 2021 (WWW '21)*, April 19–23, 2021, Ljubljana, Slovenia. ACM, New York, NY, USA, 11 pages. <https://doi.org/10.1145/3442381.3450131>

1 INTRODUCTION

The emergence of the Internet and the World Wide Web have enabled new models of collaboration on large-scale projects. Well-known examples of such peer-production systems include Wikipedia, open-source software systems such as the Linux kernel, and the collaborative world maps of Open Street Map. These systems have evolved into sophisticated organizational structures, rules, and processes that meet the challenge of managing a loose organization of contributors whose interests do not necessarily always align.

The process of maintaining a body of law in a democratic society shares many features with such peer-production systems. The work of parliaments is governed by complex rules, processes, and conventions, in order to foster compromises among competing viewpoints and priorities. How well this process works, to what

*Both authors contributed equally to this research.

This paper is published under the Creative Commons Attribution 4.0 International (CC-BY 4.0) license. Authors reserve their rights to disseminate the work on their personal and corporate Web sites with the appropriate attribution.

WWW '21, April 19–23, 2021, Ljubljana, Slovenia

© 2021 IW3C2 (International World Wide Web Conference Committee), published under Creative Commons CC-BY 4.0 License.

ACM ISBN 978-1-4503-8312-7/21/04.

<https://doi.org/10.1145/3442381.3450131>

extent it is subject to biases and to benign or undue influences is of obvious concern to citizens and to scientists alike. An exciting recent development in this regard is the adoption of *open government* principles in many countries [11, 14, 30], which aim to improve the transparency of the law-making process and the accountability of its protagonists. The European Union (EU) has been a pioneer in this: It publishes detailed records of the process by which bills are written and amended, until they finally become law. Once an initial draft of a new law has been published, parliamentarians (MEPs, for Members of the European Parliament) in one or several specialized committees examine the draft and propose amendments, consisting of one or more edits. Several edits can be in conflict if they attempt to modify the same part of the law draft. To be instituted, an edit needs to be approved by the committee in charge, and ultimately by the full plenary. The European Parliament publishes every proposed edit and its authorship, along with various other details. This makes it possible to build detailed models of the interplay between MEPs, law drafts, edits, and committees.

In our previous work [21], we curated a large-scale dataset of edits proposed by MEPs over two legislature periods (2009–2019), and we developed a predictive model for the success and failure of proposed edits. An edit can fail because the status quo is favored (*i.e.*, the edit is voted down), or because a conflicting edit is favored (*i.e.*, among mutually incompatible edits, another wins). Our model relies mostly on the structure of incompatible edits, which can be viewed as a *conflict graph* among all edits that target the same law. We learn a supervised model that endows each law with an “inertia” parameter, which captures the difficulty to amend that law, and each MEP with a “strength” parameter, which captures the influence or political savvy of the MEP. We show that the model predicts well the edit success, despite its parsimonious structure; in particular, our model does not incorporate any explicit features about the laws and edits. However, this implies an important limitation: Learning the inertia of a law requires training examples of edits success or failure *for that law*. Therefore, we were unable to make a prediction for a new draft of a law for which no edits are contained in the training set¹.

In this paper, we complement our previous dataset with additional features². Specifically, we collected explicit features for each MEP, including their party membership, country of origin, and gender. We also collected explicit features of the dossiers (law drafts) and edits, including the specific committee in charge and its type. We also collected the actual text of the edits, which enabled us to build richer models that take into account the content of the law, as well as the changes affected by each edit. The combination of these explicit features (meta-information) and of the text gives rise

¹This is reminiscent of the *cold-start problem* in recommender systems.

²Data and code publicly available on <https://github.com/indy-lab/war-of-words-2>

to models with improved predictive performance. Also, it enables us to make predictions for unseen laws. Finally, we also endow our model with a set of latent features for both laws and MEPs, which capture richer interactions between MEPs and laws than our previous model. Indeed, it would seem plausible that an MEP might be an expert in one subject matter, but less knowledgeable in another, which would bear upon their effectiveness in promoting a particular edit.

Let us briefly summarize our results. We learn a model to predict the adoption or rejection of a proposed edit. An edit can fail because it is rejected in favor of the existing version of the law (the status quo), or because another edit that it is in conflict with is accepted. In our experiments, we report the cross-entropy loss of our predictions. The main results we report assume the *new edit setting* (similar to [21]), where the edits in the data are randomly split into training, validation, and test sets. Consequently, most laws that appear in the validation and test sets have edits in the training set. We show that enriching our basic model with additional features results in a significant performance gain, and we explore their relative contributions. This exposes some rather subtle intricacies of the European Parliament’s organization and decision-making; for example, we show that the type of committee and the part of the law involved have an impact on the probability of adoption. We also explore how the latent dossier features, learned by the model, cluster into interpretable topics, and we provide some interpretation of the most predictive words and bigrams in an edit. Finally, we apply the model to the more challenging *new law setting*, where a law at test time has not been seen at training time. We show that the MEP and dossier features have a predictive value, although the performance is, not surprisingly, lower than in the new edit setting.

The remainder of the paper is structured as follows. In Section 2, we state the problem and provide a detailed description of our dataset. We describe our statistical models in Section 3. We give the results and interpretations of our experiments in Section 4. We describe related work in Section 5 and conclude in Section 6.

2 DATASET & PROBLEM STATEMENT

2.1 The EU Law-Making Process

The legislative process of the EU shares various features with the one of liberal democracies. Most laws are created through the *ordinary legislative procedure*, which works as follows. First, the European Commission (*i.e.*, the executive branch of the EU) drafts a law proposal and sends it to the European Parliament (*i.e.*, the representatives of the people in the EU). The Parliament dispatches the proposal to one of its committees (*e.g.*, for the Agriculture, for Research and Innovation, and for the Economy), whose theme is most closely related to that of the proposal. A committee is a subset of the 751 parliamentarians (MEPs, for Members of the European Parliament). For example, if the proposal is about limiting the carbon emissions in the EU, it will go to the Environment Committee.

One MEP in the committee is elected as the *rapporteur*, *i.e.*, as the person in charge of the proposal for the committee. The rapporteur and all other MEPs in the committee may propose *amendments* to the proposal, *i.e.*, modifications to parts of the law. An amendment consists of one or several *edits*, *i.e.*, a sequence of contiguous words that are added to or removed from the proposal text. These edits

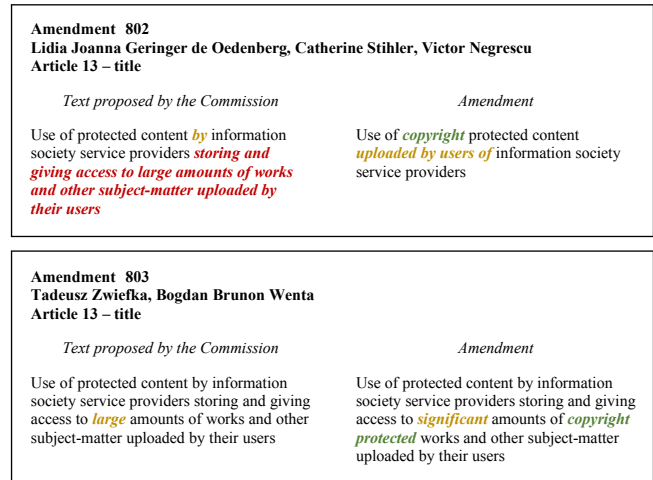


Figure 1: Example of two conflicting amendments in their raw format on the title of Article 13 of a proposal about copyrights on the Internet. (Top) Am. 802 is proposed by three MEPs and consists of three edits. (Bottom) Am. 803 is proposed by two other MEPs on the same text, and it consists of two edits. The last edit of Am. 802 (deleting the end of the title) conflicts with both edits of Am. 803. Only the first edit of Am. 803 (replacing “large” by “significant”) was accepted, and all other edits were rejected.

might conflict with other edits if they attempt to change the same part of the law but in different ways. The members of the committee vote on each edit to decide whether to include it or not in the final *report*; this decision forms the position of the Parliament on the proposal. The report is then voted by the whole Parliament: If it is accepted, it is transferred to the Council of Ministers (*i.e.*, the equivalent of a senate representing the member states of the EU). If it is rejected, the proposal is abandoned.

Optionally, MEPs in another committee may decide that their expertise is relevant to the proposal. For example, the Transportation Committee might also want to make amendments to the proposal about limiting carbon emissions in the EU. They may, therefore, send their *opinion* to the reporting committee, *i.e.*, their suggested amendments (hence, edits) to the proposal. The process is similar to that of creating a report: A rapporteur is elected to be in charge of the opinion and may, together with other MEPs in the opinion committee, propose amendments. The opinion differs from the report in that it is not voted by the whole Parliament (only the report is), and the reporting committee is free to take into account the amendments from the opinion. Amendments from the opinion committee can, however, be in conflict with amendments from the reporting committee, and MEPs from the reporting committee will also have to vote on those. Using the existing terminology, we will refer to reports and opinions as *dossiers*. We give a detailed description of the European legislative process in our previous work [21].

We show an example of conflictive edits in two amendments in Figure 1. The two amendments are proposed on Article 13 of a proposal about copyrights on the Internet. Amendment 802 is proposed by three MEPs and consists of three edits: (a) Inserting

“copyright” (in green), (b) replacing “by” by “uploaded by users of” (in yellow), and (c) deleting the end of the title after “providers” (in red). Amendment 803 is proposed by two other MEPs and consists of two edits: (d) Replacing “large” by “significant” (in yellow) and (e) inserting “copyright protected” (in green). There are two conflicts in this amendment: Edit (c) of the first amendment is in conflict with Edit (d), and it is also in conflict with Edit (e). All these edits are also implicitly in conflict with the original text proposed by the European Commission. Out of these five edits, only Edit (d) was accepted. All other edits were rejected, *i.e.*, the status quo was voted and the text proposed by the Commission was maintained.

2.2 Explicit Features

Our original dataset contained the following metadata: (a) The author(s) of an amendment (b), the dossier that is amended, and (c) the rapporteur for this dossier. We complement this dataset by extracting explicit (meta) features of the MEPs, the edits, and the dossiers, as well as text features. For each MEP, we collect their nationality (one of 28), their EU political group (one of 9), and their gender. A political group clusters national parties that share similar political ideologies. For each edit, we identify whether it is an insertion, a deletion, or a replacement of some words in the proposal, and we compute its length. We also collect information about where in the law the edit was proposed: in an article (in the body of the proposal), in a recital (in the preamble of the proposal), in an annex, or in other more specific but less frequent parts of a law. We determine whether an edit in a reporting committee comes from an opinion committee (in which case it is an “outsider”). Finally, we note whether an edit comes with an optional justification. For each dossier, we identify its type (report or opinion) and the committee that is in charge. We also note if the proposal is a regulation (legally binding for all member states of the EU), a directive (sets general goals that member states can implement however they want), or a decision (binding to one member state or company only). We describe these explicit features in Table 1.

In total, we collect 449 493 edits from 237 177 amendments in the European Parliament during the 7th and the 8th legislature periods³ (referred to as EP7 and EP8), between 2009 and 2019 (each period lasts 5 years). After gathering the edits according to the conflicts, we obtain 267 451 conflicts for both EP7 and EP8, covering 1889 dossiers. We summarize this dataset in Table 2.

2.3 Text Features

We further augment the dataset by collecting text features of the edit itself. It is reasonable to expect that certain words and phrases are predictive of the success of an edit. We extract the deleted words w_- from the proposal and the inserted words w_+ from the amendment. In Figure 1, for example, Edit (b) of Amendment 802 has $w_- = \text{“by”}$ and $w_+ = \text{“uploaded by users of”}$. We also consider the context of an edit by extracting the original text of the amended article, surrounding the location of the edit. For Edit (b) in Amendment 802, the context consists of the two portions of text “Use of protected content” and “information society...their users”. Finally, we

³We do not collect data from EP9 (2019 – 2023), as the amount of published data is too small at this time: The legislature period started in Fall 2019 and the Parliament’s activities were slowed down due to the COVID-19 crisis in Spring 2020.

Table 1: List of features for MEPs and edits.

Category	Feature	Type [Values]
MEP	Nationality	Categorical [28]
	Political group	Categorical [8 or 9]
	Gender	Categorical [2]
Edit	Rapporteur	Binary
	Edit type	Categorical [3]
	Log-length (+)	Numerical [$\mathbb{R}_{\geq 0}$]
	Log-length (-)	Numerical [$\mathbb{R}_{\geq 0}$]
	Article type	Categorical [7]
	Outsider committee	Binary
Dossier	Justification	Binary
	Type	Categorical [2]
	Committee	Categorical [35]
	Legal act	Categorical [3]

also extract the title of the law proposal; we will use it as a text feature of the dossier. For Amendments 802 and 803, the title is “Copyright in the Digital Single Market”. We map all words to lower case, and we replace digits in the title by the letter “D”, as there are many reference numbers that are unlikely to be useful for our task.

We give some statistics of the distribution of the length of the deleted text w_- , the inserted text w_+ , the context, and the title in Table 3. We report the lower quartile Q_1 and the upper quartile Q_3 , as well as the median. About half of the inserted and deleted texts are short (7 words or less), but the distribution of lengths has a long tail, as shown by the larger values of the upper quartile Q_3 . The context provides large portions of text (the median is at 42 for EP7 and 49 for EP8), which will be useful for making predictions. In Section 3, we describe how we incorporate the explicit features and the text features into our models.

2.4 Problem Statement

We build a model that predicts the vote outcome of edits that will form the reports and the opinions. Formally, we take a supervised approach to solve the following prediction problem: Let $C = \{a, b, \dots\}$ be a set of conflictive edits proposed on a dossier i , for which we have observed other edits. We want to predict which of the conflictive edits in C or the status quo of the proposal for dossier i will be accepted within the committee. This task differs from multinomial classification as the number of classes varies for each data point: If an edit a is in conflict only with the original text proposed by the Commission, then $|C| = 1$. If several edits $a, b, \dots \in C$ are in conflict against each other, then $|C| > 1$.

According to Rule 180 of the Rules of Procedure of the European Parliament [26], the committee sets a deadline by which MEPs must propose amendments to a dossier. The voting takes place after this time. Hence, at the time of voting, an edit is expected to confront all alternatives: If edits a , b , and c are in conflict, the MEPs vote on all three of them and the status quo to select only one outcome.

3 STATISTICAL MODELS

To better introduce our models, we first define the baselines against which we will compare our results in Section 4. In particular, we

Table 2: Descriptive statistics of our extended dataset.

Period	# am.	# edits	# conflicts	# MEPs	# dossiers	% accepted	% inserted	% deleted	% replaced
EP7 (2009–2014)	108 292	200 407	126 417	761	1 089	37.7%	37.8%	22.0%	40.2%
EP8 (2014–2019)	128 885	249 086	141 034	791	800	25.7%	37.9%	22.4%	39.7%

Table 3: Distribution of text lengths in number of words.

Legislature	Type	Q_1	Median	Q_3
EP7	Insertion w_+	2	7	20
	Deletion w_-	2	6	26
	Context	15	42	79
	Title	6	12	19
EP8	Insertion w_+	2	6	17
	Deletion w_-	2	6	28
	Context	20	49	93
	Title	6	10	22

recall the WAR OF WORDS model, as introduced in our previous work, and we adopt the same terminology for consistency. For each baseline and for our models, we assume a set of K conflicting edits $C = \{a, b, \dots\}$ proposed on dossier i , for which we want to model the probability that an edit $a \in C$ is accepted over edits b, \dots on this dossier. We denote this probability by $p(a >_i C - \{a\})$, and we denote the probability that the status quo wins, *i.e.*, that the original text proposed by the Commission is kept, by $p(i > C) = 1 - \sum_{a \in C} p(a >_i C - \{a\})$.

3.1 Baselines

Naive Classifier. The *naive classifier* predicts a uniform probability for each outcome, *i.e.*, for each of the conflicting edits or the status quo to win, as

$$p(a >_i C - \{a\}) = p(i > C) = \frac{1}{K+1}.$$

Random Classifier. The *random classifier* learns the prior probability $p^{(K)}$ that the status quo wins for each conflict size $|C| = K$, and it predicts

$$p(i > C) = p^{(K)}.$$

It predicts uniformly each of the edits to win as

$$p(a >_i C - \{a\}) = \frac{1 - p^{(K)}}{K}.$$

WAR OF WORDS. The WoW model encodes (a) the collaboration between MEPs who co-sponsor an edit and (b) the conflicts between edits as a discrete-choice model reminiscent of the Bradley-Terry model [3] and the Rasch model [27]. It models the probability that an edit a is accepted over edits b, \dots on dossier i as

$$p(a >_i C - \{a\}) = \frac{\exp(s_a)}{\sum_{c \in C} \exp(s_c) + \exp(d_i + b)}, \quad (1)$$

where $s_a = \sum_{u \in \mathcal{A}_a} s_u$ is the cumulated *skill* of all authors \mathcal{A}_a of edit a , $d_i \in \mathbf{R}$ is the *difficulty* of dossier i , and $b \in \mathbf{R}$ is a global bias parameter. The skill parameters s_u of the MEPs can be interpreted

as a measure of their influence, and the difficulty parameters d_i of the dossiers can be interpreted a measure of their controversy.

3.2 Enriched Models

Explicit Features. We extend the WoW model by augmenting it with explicit features of the MEPs (e.g., nationality), the edits (e.g., length of inserted text), and the dossiers (e.g., report or opinion), as described in Table 1. From (1), we replace the skill parameters s_a with the inner product between a feature vector $\mathbf{s}_a \in \mathbf{R}^{M_E}$ of M_E features of edit a and the associated parameter vector $\mathbf{w}_E \in \mathbf{R}^{M_E}$. We also replace the difficulty parameter d_i by the product of a feature vector $\mathbf{d}_i \in \mathbf{R}^{M_D}$ of M_D features of dossier i and its associated parameter vector $\mathbf{w}_D \in \mathbf{R}^{M_D}$. We then have

$$p(a >_i C - \{a\}) = \frac{\exp(\mathbf{s}_a^\top \mathbf{w}_E)}{\sum_{c \in C} \exp(\mathbf{s}_c^\top \mathbf{w}_E) + \exp(\mathbf{d}_i^\top \mathbf{w}_D + b)}. \quad (2)$$

We refer to this model as WoW(*Explicit*) (or WoW(X), for conciseness). In (1), the feature vector \mathbf{s}_a is the indicator of the authors of an edit a : Its entries s_u are 1 for all $u \in \mathcal{A}_a$ and 0 otherwise. Similarly, the feature vector \mathbf{d}_i is the indicator of dossier i . In (2), the feature vectors \mathbf{s}_a and \mathbf{d}_i represent features related to MEPs, edits, and dossiers derived from our dataset.

Latent Features. Consider the simple case of an MEP u proposing an edit on dossier i , and suppose that this edit conflicts with another edit, proposed by MEP v . From (1), let $p(u >_i v)$ be the probability that, for dossier i , the edit proposed by MEP u is accepted over the edit proposed by MEP v . The assumption made in the WoW model is strong: It posits that if MEP u is more influential than MEP v , then, all other things being equal, $p(u >_i v) > p(v >_i u)$ for all dossiers i . This assumption is not always realistic: Dossiers span a vast amount of different topics, and the MEPs have their own specializations and interests. For example, an MEP familiar with fisheries might not be knowledgeable about research and academia.

In order to capture these dependencies, we incorporate a bilinear term into the WoW model. We assign a vector $\mathbf{x}_u \in \mathbf{R}^L$ to each MEP u , and a vector $\mathbf{y}_i \in \mathbf{R}^L$ to each dossier i , for some dimensionality $L > 0$. We then rewrite (1) as

$$p(a >_i C - \{a\}) = \frac{\exp(s_a + \mathbf{x}_a^\top \mathbf{y}_i)}{\sum_{c \in C} \exp(s_c + \mathbf{x}_c^\top \mathbf{y}_i) + \exp(d_i + b)}, \quad (3)$$

where $\mathbf{x}_a = \sum_{u \in \mathcal{A}_a} \mathbf{x}_u$ is the sum of the latent features \mathbf{x}_u of each author u of edit a . We refer to this model as the WoW(*Latent*) model (or WoW(L)). The latent vectors \mathbf{x}_u and \mathbf{y}_i can be viewed as the embeddings of MEP u and of dossier i in a Euclidean latent space. Informally, the probability $p(a >_i C - \{a\})$ increases when the MEP embedding \mathbf{x}_a is co-linear with the dossier embedding \mathbf{y}_i in the latent space. It decreases when the two vectors point in opposite directions. Furthermore, vector \mathbf{x}_u can be interpreted as the set of

skills of MEP u . Similarly, \mathbf{y}_i can be interpreted as the set of skills required to edit dossier i .

Text Features. The features described so far ignore the text content of the edit itself. It is reasonable to expect that the presence of certain words or phrases in the original or amended text of an edit, and in the title of the dossier, are predictive of the success of the edit. Hence, we incorporate text features to the WoW model by rewriting (1) as

$$p(a >_i C - \{a\}) = \frac{\exp(s_a + \mathbf{r}_a^\top \mathbf{w}_T)}{\sum_{c \in C} \exp(s_c + \mathbf{r}_c^\top \mathbf{w}_T) + \exp(d_i + \mathbf{r}_i^\top \mathbf{w}_{T'} + b)}, \quad (4)$$

where $\mathbf{r}_a \in \mathbf{R}^D$, $\mathbf{r}_i \in \mathbf{R}^{D'}$ are, respectively, representations of the text of the edit a and the title of dossier i , and $\mathbf{w}_T \in \mathbf{R}^D$, $\mathbf{w}_{T'} \in \mathbf{R}^{D'}$ are, respectively, the associated parameter vectors. We refer to this model as the WoW(*Text*) model (or WoW(*T*)).

We explore different ways of learning the representations \mathbf{r}_a and \mathbf{r}_i from (a) pre-trained word embeddings and (2) by training embeddings on our dataset. With pre-trained embeddings, \mathbf{r}_a is the concatenation of three vectors that are the representations of the deleted text, inserted text, and the context of the edit, as explained in Section 2. Each of these vectors are the averages of the pre-trained word embeddings of the words in these parts of the text, and \mathbf{r}_i is the average of the pre-trained embeddings of the words in the title of dossier i . We use two sets of pre-trained embeddings trained with the word2vec algorithm [24]: (a) 300-dimensional embeddings trained on Google News [10] and (b) 200-dimensional Law2Vec embeddings trained on legal texts of the EU, the US, the UK, Canada, and Japan[5].

We also learn embeddings from our dataset by using the supervised fastText model for text classification [17]. In the simplest version of this model, a D -dimensional embedding is learned for each word (and n -grams) in a dataset. A piece of text is then classified with a softmax layer by representing it as the average of the word embeddings. We use the learned word and bigram embeddings to construct \mathbf{r}_a and \mathbf{r}_i .

The original fastText model is defined, however, for classification of homogeneous pieces of text into a fixed set of classes. This does not directly apply to our problem, as (a) the text features for the edit are of three types (deleted text, inserted text, and context) and (b) the size of a conflict $|C| = K$ varies from a data point to another. We solve the first problem by prepending tags (, <ins>, and <con>) to each word to enable the model to learn separate embeddings for the same word in different types of text feature. We solve the second problem by training the embeddings on a binary classification task of edit acceptance (based only on the text), and by using the embeddings learned on this ad-hoc task into the WoW models. We learn the embeddings for the words in the title by training a different fastText model to predict the acceptance of an edit from the title only. This is equivalent to predicting the probability of acceptance of the status quo for each dossier, given its title. For our experiments in Section 4, we use the fastText embeddings rather than pre-trained embeddings, because the former performed better on the ad-hoc binary classification task.

Table 4: Variations of our model by combination of features (explicit, latent, and text features).

Model	Equation	Explicit	Latent	Text
WoW	(1)	–	–	–
WoW(<i>Explicit</i>)	(2)	✓	–	–
WoW(<i>Latent</i>)	(3)	–	✓	–
WoW(<i>Text</i>)	(4)	–	–	✓
WoW(<i>XL</i>)	(2) & (3)	✓	✓	–
WoW(<i>XT</i>)	(2) & (4)	✓	–	✓
WoW(<i>LT</i>)	(3) & (4)	–	✓	✓
WoW(<i>XLT</i>)	(2), (3) & (4)	✓	✓	✓

Hybrid Models. We combine WoW(*Explicit*), WoW(*Latent*), and WoW(*Text*) together to obtain hybrid models with different components. This helps us understand the contribution of each type of features to the performance, in Section 4. We summarize all the possible combinations in Table 4, and we sort them by increasing levels of complexity. The WoW model has no features at all and will serve as a baseline. The WoW(*XLT*) combines explicit, latent, and text features together, and it has the highest complexity.

3.3 Learning the Parameters

Each observation n is a triplet (C_n, i_n, l_n) of (a) a set of conflicting edits C_n with $|C_n| = K_n > 0$, (b) a dossier i_n on which the edits are proposed, and (c) a label $l_n \in C_k \cup \{i_n\}$ indicating which of the K_n edits or the status quo is accepted. We assume that the triplets are independent. Given a dataset of N triplets $\mathcal{D} = \{(C_n, i_n, l_n) \mid n = 1, \dots, N\}$ and given a vector θ of all the parameters in our model, we learn θ by minimizing their negative log-likelihood under \mathcal{D}

$$-\ell(\theta; \mathcal{D}) = \sum_{n=1}^N \sum_{a \in C_n} \left[\mathbf{1}_{\{l_n=a\}} \log p(a >_{i_n} C_n - \{a\}) + \mathbf{1}_{\{l_n=i_n\}} \log p(i_n > C_n) \right],$$

where $p(a >_{i_k} C_k - \{a\})$ and $p(i_k > C_k)$ depend on θ . In order to avoid overfitting, we add regularization to the negative log-likelihood. We pre-process our dataset by keeping only the dossiers for which more than 10 edits have been proposed and only the MEPs who have proposed more than 10 edits. Hence, we obtain a dataset of $N = 125733$ data points for EP7 and $N = 140763$ data points for EP8. In the WoW(*Explicit*) and the WoW(*Text*) models, the log-likelihood is convex, and we find optimal parameters by using an off-the-shelf convex optimizer (L-BFGS-B [4]). In the WoW(*Latent*) model, the bi-linear term breaks the convexity, and we can no longer ensure that we will find parameters that are global optimizers. In practice, by using a stochastic gradient descent algorithm (Adagrad [9]), we are still able to find good model parameters without convergence issues.

4 RESULTS

4.1 Experimental Setting

We report the cross-entropy loss to evaluate the baselines and our models. Let (C_n, i_n, l_n) be an observation. We compute

$$\ell_n = \begin{cases} -\log p(l_n > i_n | C_n - \{l_n\}) & \text{if } l_n \in C_n, \\ -\log p(i_n > C_n) & \text{if } l_n = i_n. \end{cases} \quad (5)$$

We report the average value for all N points in our test set as $\ell = \frac{1}{N} \sum_n \ell_n$. We randomize our dataset and we split it into 80% for training, 10% for validation, and 10% for the final evaluation. Note that an edit c is involved in several conflicts. For example, in Figure 1, edit c is involved in two conflicts: $C_1 = \{c, d\}$ and $C_2 = \{c, e\}$. Hence, we assign conflicts to each set so that an edit is present in exactly one set. We combine both the training and the validation sets to fit our model before evaluating it on the test set. We set the number of latent dimensions L and the regularizers, and we choose the best word embeddings, by held-out validation. This results in fastText of dimension $D = D' = 10$, with bigrams.

4.2 Predictive Performance

We show in Figure 2 the overall performance of all variations of our model (with and without explicit, latent, and text features) over EP7 and EP8, and we compare them against the naive and the random predictors, as well as against the WoW model. All our models outperform the baselines, and WoW(XLT) outperforms all other models. Including explicit features improves the performance of the predictions in terms of the cross entropy by 7% for EP7 and 6% for EP8 over the simpler WoW model. On EP7, WoW(L) improves the performance by 12% and WoW(T) by 7%, whereas for EP8 the difference between the two models is smaller (10% increase for WoW(L) and 8% for WoW(T)). Hence, the text features provide a greater improvement for EP8 than for EP7, while the latent features provide a greater improvement for EP7 than for EP8. The difference between WoW(XL) and WoW(L) (0.010 for EP7 and 0.013 for EP8) is less than the difference between WoW(XT) and WoW(T) (0.034 for EP7 and 0.035 for EP8), as the latent features absorb the effects of the explicit features more than the text features do. Finally, combining the text and latent features provides high performance, but further combining them with explicit features leads to the best performance.

4.3 Interpretation of Explicit Features

To understand the contribution of the explicit features to the predictive performance, we show in Figure 3 the decrease in cross-entropy loss of WoW(MEP) (all MEP features but the rapporteur feature), WoW(Rapporteur) (rapporteur feature only), WoW(Edit), and WoW(Dossier) over WoW. The dossier features contribute virtually nothing to the predictive performance (the difference is at the fourth decimal point). Similarly, for EP7, the nationality, political group, and gender features of WoW(MEP) contribute very little. For EP8, these features improve the performance, but not as much as the edit features. This suggests that these features have limited influence on the predictions. Nationalities and political groups have been qualitatively analyzed in the literature in the context of their influence on MEPs' voting behaviour [6, 13, 22, 25]. To the best of

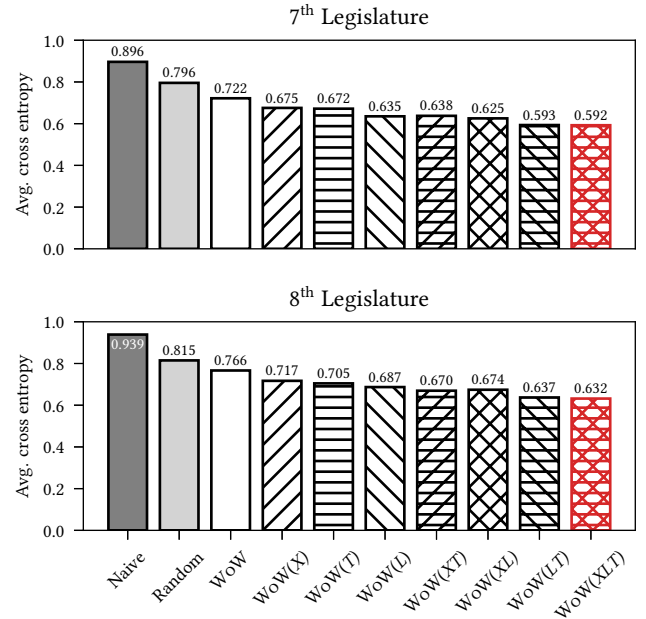


Figure 2: Average cross-entropy loss of the baselines and our models. Combining the explicit, latent, and text features help obtain the best performance.

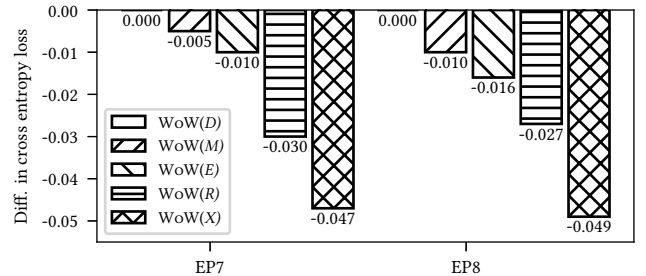


Figure 3: Difference in cross-entropy loss over WoW of different models. The rapporteur feature and the edit features contribute more to the predictive performance than the MEP and dossier features.

our knowledge, there is no analysis of their effect on the amending process. Interestingly, for EP7, combining all features into the WoW(X) model leads to a performance boost that is greater than the sum of each individual feature groups.

To get insights into the dynamics of the legislative process, we interpret the values of the parameters of WoW(XLT) trained on the full dataset for EP8 (combining training, validation, and test data). Let $w_f \in \mathbb{R}$ be the value of the parameter associated with feature f . The rapporteur feature r of WoW(Rapp.) provides a greater decrease in loss. This *rapporteur advantage* complements the findings of Costello and Thomson [7], conducted by interviewing key informants over EP5 (1999-2004) and EP6 (2004-2009). They show that the rapporteur, with their particular role, has some influence on the legislative process, albeit constrained. We note that, according

to our model, the rapporteur advantage has slightly increased in EP8 ($w_r = 1.19$) compared to EP7 ($w_r = 1.12$).

These explicit features enable us to explain what contributes to the success of an edit. We report here (and in subsequent sections) the results for EP8 only. All other things being equal, a female ($w_{\text{fem}} = -0.02 > -0.04 = w_{\text{mal}}$) MEP from Latvia and whose party belongs to the group of the European People’s Party (center-right) has the highest chance to see her edit accepted. This edit has even higher chances if it inserts ($w_{\text{ins}} = -0.03 > w_{\text{del}} = -0.13 > w_{\text{rep}} = -0.22$) a short portion of text (the feature associated with both insertion and deletion length is negative) in a part of the law that is not its body or its preamble (w_{art} , w_{rec} and w_{para} have the lowest value among the seven article types). Adding a justification also increases the probability of an edit being accepted ($w_{\text{jus}} = 0.08$), as well as edits from the opinion committee (referred to as the “outsider committee” feature in Table 1, $w_{\text{out}} = 0.16$).

For the dossier features, our model learns that it is harder to make edits on reports, as compared to opinions ($w_{\text{rep}} = 0.33 > -0.26 = w_{\text{opi}}$). As explained in Section 2, reports are voted by the whole Parliament. Therefore, they have a greater influence on the final law, and we expect that MEPs make it more difficult for competing edits to be accepted in reports. Finally, our model also learns that it is harder to make edits for decisions and directives, as compared to regulations ($w_{\text{dec}} = 0.25 > w_{\text{dir}} = 0.12 > w_{\text{reg}} = 0.10$).

4.4 Interpretation of Text Features

In Figure 2, we observe that the text features contribute significantly to improving the performance. We use the learned parameter vectors w_T and $w_{T'}$ of WoW(XLT) to identify words and bigrams that have the most predictive power. First, we rank the words and bigrams of the edit text, according to the dot product of their embeddings with w_T . The top- k terms (having a positive dot product) contribute the most towards acceptance of the edit, whereas the bottom- k terms (having a negative dot product) contribute most towards rejection of the edit. The opposite holds for the terms of the title and their dot product with $w_{T'}$.

We look at the top 50 terms for each feature and prediction outcome and find some interesting patterns among these terms, although not all of them are easy to interpret. Note that we have more than 10 000 unique terms for the edit text and more than 1 000 unique terms for the title, hence we consider only the most predictive terms near the ends of the ranking. A list of the top-50 terms for each feature and prediction outcome is reported in Appendix A.

One of the bigrams that, when deleted, is predictive of acceptance is *any other*, which is commonly used to widen the scope of the law (as in “contractual or any other duty”). Interestingly, the bigrams *human rights* and *data protection* are also predictive of acceptance when deleted. The word *should*, which is used to add recommendations, is predictive of acceptance when inserted, while adding *must*, which is used for obligations, is predictive of rejection. We see that *best* is predictive of acceptance, which is commonly used to make a requirement stronger (as in “best available scientific evidence”, “best possible way”). Adding *positive* and *positive impact* predicts acceptance, whereas adding *negative* predicts rejection. Adding the word *inserted*, which commonly refers to inserting new

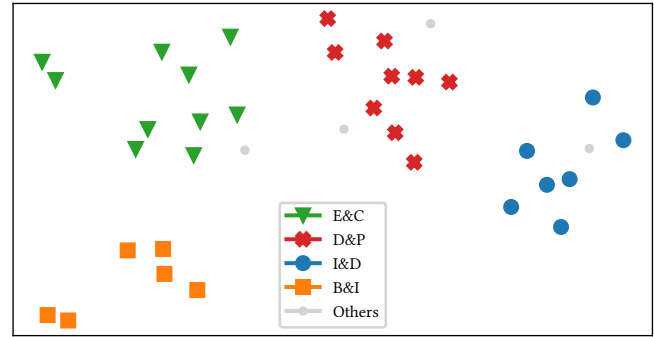


Figure 4: Visualization with t-SNE of the top-10 and bottom-10 dossiers on the first two principal components in EP8. There are four clusters: Environment and Communication (E&C), Defense and Protection (D&P), Investment and Development (I&D), and Business and Innovation (B&I).

articles in existing laws, is predictive of acceptance, whereas *deleted* is predictive of rejection.

Considering the words in the context, we see that *firearms*, *re-settlement*, *terrorist* and *fingerprints* are predictive of rejection. This could be because the laws related to these topics are controversial, hence many edits are rejected due to conflicts. For the words in the title, we see that *customs*, *community*, *financial*, *fisheries*, and *general budget* are predictive of acceptance, whereas *market*, *framework*, *structural reform*, *emission*, and *greenhouse gas* are predictive of rejection. This suggests the relative ease or difficulty of editing laws related to these topics, and it correlates well with the values of the difficulty parameters d_i : The top-50 dossiers with the highest difficulty parameters contain high-controversy dossiers about establishing frameworks for the screening of foreign investments and vast public investment programs (InvestEU and Horizon Europe), as well as regulation of the financial market, copyright in the digital market, and carbon-emission reduction. The bottom-50 dossiers with the lowest difficulty parameters contain low-controversy dossiers about cohesion within the EU, financial rules, fisheries, and the community code on visas.

4.5 Interpretation of Latent Features

The latent features improve the predictions overall and help capture the complex dynamics of the legislative process. The best number of latent dimensions is $L = 20$ for the models including latent features. In order to interpret the latent features, we gather the latent vectors y_i learned by WoW(XLT) into a matrix $Y = [y_i]$. We apply principal component analysis and keep the top-10 and bottom-10 dossiers from each of the first two principal components in EP8. We use t-SNE [23] to represent these forty dossiers in a two-dimensional space, and we show the projection in Figure 4.

We distinguish four clusters. The cluster at the top-left contains dossiers about fuel quality, renewable energy, trade of animals, and sustainable investments. It also contains dossiers about electronic communications, the processing of personal data, and sharing public information. We interpret this cluster as *environment and*

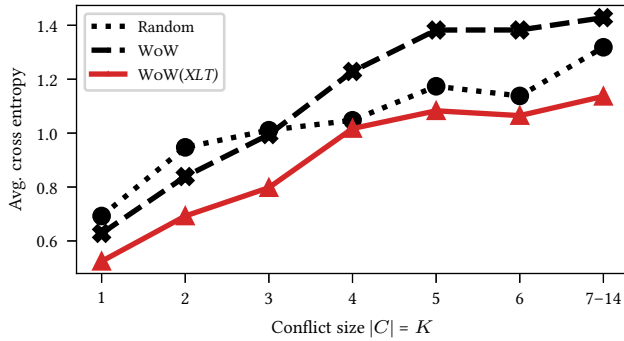


Figure 5: Average cross-entropy loss per conflict size $|C| = K$. The loss of the WoW(XLT) model increases less rapidly than the loss of the baselines.

communication, and we highlight with green triangles the corresponding dossiers. The cluster at the top-center contains dossiers about the establishment of defense funds, the prosecution of criminal offenses, and the identification of criminals between member states. It also contains dossiers about the protection of workers, businesses, refugees, internal markets, and cultural goods. We interpret this cluster as *defense and protection* (red crosses). The cluster at the top-right contains dossiers about vast investment and development programmes, finance, and the development of internal markets. We interpret this cluster as *investment and development* (blue dots). Finally, the cluster at the bottom-left contains dossiers about economic competitiveness and innovation, as well as frameworks for business development and the funding of start-up companies. We interpret this cluster as *business and innovation* (orange squares).

4.6 Error Analysis by Conflict Size

We explore how the WoW(XLT) model performs on conflict of different sizes in the test set for EP8 (we observe a similar behaviour on EP7). We bin the conflict size so that there are at least 100 data points in each bin. The distribution of conflict size is exponentially decreasing: There are 8462 conflicts of size 1 (i.e., an edit is in conflict with the status quo only), 3063 conflicts of size 2 (i.e., two edits are in conflict, as well as with the status quo), and 140 conflicts of size 7 and more. We compare the average cross entropy of the WoW(XLT) model with that of the random predictor and that of the WoW model. In Figure 5, we see that while the loss generally increases with conflict size for all three models, it increases less rapidly for the WoW(XLT) model than for the WoW model. This suggests that the explicit, latent, and text features enable the model to exploit the increasing complexity of data points to make more accurate predictions. We also see that for conflicts of size 4 and higher, the WoW model performs worse than the random predictor, but the WoW(XLT) model is able to outperform it.

4.7 Solving the Cold-Start Problem

We explore how to solve the cold-start problem by defining a second predictive problem: Given a dossier i for which we have never seen an edit, and given a conflict $C = \{a, b, \dots\}$, we want to predict which of the edits or the status quo wins. We order the dossiers by the date a

Table 5: Average cross entropy of the baselines and our model on predicting new, unseen dossiers.

Type	Model	Avg. cross entropy
Baseline	Naive	0.947
	Random	0.800
	WoW	0.873
Ours	WoW(Explicit)	0.784
	WoW(Text)	0.839
	WoW(XT)	0.759

committee received a proposal, and we use the dossiers that contain the first 80% of the conflicts as a training set. We use the next 10% as validation set, and we keep the last 10% aside as test set. We ensure that no edits in the training set leak into the validation and test sets. This scenario is more realistic because we make predictions about new dossiers that the model has never observed before.

We report, in Table 5, the results for WoW(Explicit), WoW(Text), and WoW(XT), together with the baselines. The latent features cannot be used for this task, as the dossier embeddings \mathbf{y}_i are unavailable for new dossiers. For our models, the difficulty parameter d_i is set to the average difficulty learned in the training set. The random predictor, which learns the prior probability of the status quo winning for each conflict size, performs the best out of all the baselines, and it outperforms WoW(Text). Our approach outperforms only the random predictor when including explicit features. This suggests that the dossier features help us make more accurate predictions by learning parameter values for the type of dossier, its legal act, and its committee in charge. In this case, adding text features further boosts the performance.

The overall performance, however, is mixed: The improvement of WoW(XT) over the random predictor is rather small. One possible explanation is that the legislative process might be non-stationary. Hence, our model overfits on the training set, which is very different from the test set. The task is also unfair to our model, as in a real setting, predictions would be made for the next dossier only. In the current setting, we make predictions for all future dossiers. We keep further investigations of this aspect for future work.

5 RELATED WORK

This work extends our previous dataset [21] by including meta-data features from the MEPs, the edits, and the dossiers, and text features from the edits and the title of the proposals. We augment our previous model by including these explicit features and text features into the WoW model. To strengthen the model, we also borrow from collaborative filtering techniques in the recommender systems literature. Similarly to matrix factorization techniques [18] that learn latent features for users and items to make recommendations, our model learns latent features for the MEPs and dossiers to predict edit outcomes. We show that these latent features improve the predictive performance of our model by capturing bi-linear interactions between the MEPs and the dossiers.

Amendment analysis in the European Parliament has been studied by the political science community on datasets of small size [2, 19, 20, 29]. Predicting edits on collaborative corpora of documents

has been studied in the context of peer-production systems, such as Wikipedia [1, 8, 28] and the Linux kernel [15, 31]. In this work, we combine the two by taking a peer-production viewpoint on the law-making process, and by proposing a model of the acceptance of the legislative edits. Our approach generalizes to any peer-production system in which (meta) features of the users and items can be extracted and in which edits can be in conflict with one another.

We use the text of the edits and dossiers as features for classification. Text classification is a well-studied problem in natural language processing. A simple baseline is to apply linear classifiers to term-frequency inverse document-frequency (TF-IDF) vectors [16]. However, these models do not capture the synonymy relation between words, hence suffer from poor generalization. Models based on neural networks show better performance on this task [33]. They tend, however, to require larger datasets, and the features they learn are harder to interpret. The fastText model [17] bridges the gap between the two: It learns embeddings from linear models. We adapt this approach to our problem of edit classification, as edits are inhomogeneous pieces of text. Edit modelling has been studied using neural models [12, 32] that suffer from the aforementioned issues of dataset size and interpretability. In the WoW models, we combine text features and non-text features to take into account the dynamics of the legislative process. Legal texts also have features and structures that set them apart from other domains. For example, the word "should" has a strong legal significance, whereas it is commonly removed as a stop word.

6 CONCLUSION

In this paper, we extended our previous work on predicting legislative edits, where we considered influence parameters of the MEPs, controversy parameters of the dossiers, and the rapporteur advantage. We complemented our dataset with (a) additional explicit features of the edits, of the MEPs, and of the dossiers, (b) latent features of the MEPs and dossiers, and (c) text features of the edits and dossiers. Each of the three classes of additional features improve the performance significantly, and the best performance is achieved by combining all features. We interpreted the values of the learned parameters to gain insights into the legislative process. We provided interpretation of all explicit features to characterize what makes the success of an edit more likely. We have shown that the latent features capture the representation of MEPs and dossiers in an ideological space. We have analyzed the words and bigrams in different parts of an edit and a dossier in terms of their influence on the acceptance probability. We have also analyzed the performance of our model on subsets of the test set based on conflict size, and we have shown that our best model can leverage the features of the data to make more accurate predictions on conflicts of higher size than other baselines. Finally, we have described how to use our model for predicting edits made on new, unseen dossiers.

Ethical Considerations. One anonymous reviewer expressed concerns regarding the use of machine learning for making decisions in law making, and whether our findings in Section 4 could help adversarial attacks. However, we do not propose to rely on our models for making decisions, such as whether an edit should be accepted or not. Our goal is to understand the factors correlated with the acceptance of edits, and thereby gain insights into the

law-making processes. These correlations do not imply a causal relationship that would benefit potential adversarial attackers.

Applications and Broader Impact. We believe that approaches such as ours are helpful to political scientists, journalists and transparency observers, and to the general public: First, it could be useful in validating theoretical hypotheses using large-scale datasets and advanced computational methods. Second, it could help uncover lesser-known facts, such as controversial dossiers that slipped under the radar. Finally, the greater transparency that results from these insights can enhance trust in public institutions and strengthen democratic processes.

Future Work. First, we currently use pre-trained word embeddings and embeddings trained on an ad-hoc binary classification task. We plan to explore how to learn text embeddings in an end-to-end manner using the conflictive structure of the WoW model. Second, as shown in Section 4.7, our model has only limited predictive power on edits made on future dossiers. We plan to further explore how to exploit the temporality of the data and how to develop a dynamical model able to take into account the non-stationarity of the law-making process. Finally, we plan to explore more complex models of textual edits, such as considering pairs of words that are inserted and deleted and longer-range word order.

ACKNOWLEDGMENTS

We thank Holly Cogliati-Bauereis and the anonymous reviewers for careful proof-reading and constructive feedback.

REFERENCES

- [1] B. Thomas Adler and Luca de Alfaro. 2007. A Content-Driven Reputation System for the Wikipedia. In *Proceedings of WWW'07*. Banff, AB, Canada.
- [2] Inger Baller. 2017. Specialists, party members, or national representatives: Patterns in co-sponsorship of amendments in the European Parliament. *European Union Politics* 18, 3 (2017), 469–490.
- [3] Ralph Allan Bradley and Milton E. Terry. 1952. Rank Analysis of Incomplete Block Designs: I. The Method of Paired Comparisons. *Biometrika* 39, 3/4 (1952), 324–345.
- [4] Richard H Byrd, Peihuang Lu, Jorge Nocedal, and Ciyou Zhu. 1995. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing* 16, 5 (1995), 1190–1208.
- [5] Ilias Chalkidis and Dimitrios Kampas. 2019. Deep learning in law: early adaptation and legal word embeddings trained on large corpora. *Artificial Intelligence and Law* 27, 2 (2019), 171–198.
- [6] Emanuel Emil Coman. 2009. Reassessing the influence of party groups on individual members of the European Parliament. *West European Politics* 32, 6 (2009), 1099–1117.
- [7] Rory Costello and Robert Thomson. 2010. The policy impact of leadership in committees: Rapporteurs' influence on the European Parliament's opinions. *European Union Politics* 11, 2 (2010), 219–240.
- [8] Gregory Druck, Gerome Miklau, and Andrew McCallum. 2008. Learning to Predict the Quality of Contributions to Wikipedia. In *Proceedings of WikiAI 2008*. Chicago, IL, USA.
- [9] John Duchi, Elad Hazan, and Yoram Singer. 2011. Adaptive subgradient methods for online learning and stochastic optimization. *Journal of Machine Learning Research* 12, Jul (2011), 2121–2159.
- [10] Google. 2013. word2vec. <https://code.google.com/archive/p/word2vec/> Accessed: 2020-10-19.
- [11] Swiss Government. 2021. Swiss Open Government Data. <https://opendata.swiss/en/> Accessed: 2021-02-14.
- [12] Kelvin Guu, Tatsunori B Hashimoto, Yonatan Oren, and Percy Liang. 2018. Generating sentences by editing prototypes. *Transactions of the Association for Computational Linguistics* 6 (2018), 437–450.
- [13] Simon Hix. 2002. Parliamentary behavior with two principals: preferences, parties, and voting in the European Parliament. *American Journal of Political Science* (2002), 688–698.
- [14] President Barack Obama's White House. 2018. Open Government Initiative. <https://obamawhitehouse.archives.gov/open> Accessed: 2020-10-19.

- [15] Yujuan Jiang, Bram Adams, and Daniel M. German. 2013. Will My Patch Make It? And How Fast? Case Study on the Linux Kernel. In *Proceedings of MSR 2013*. San Francisco, CA, USA.
- [16] Thorsten Joachims. 1998. Text categorization with support vector machines: Learning with many relevant features. In *European conference on machine learning*. Springer, 137–142.
- [17] Armand Joulin, Édouard Grave, Piotr Bojanowski, and Tomáš Mikolov. 2017. Bag of Tricks for Efficient Text Classification. In *Proceedings of ACL 2017*.
- [18] Yehuda Koren, Robert Bell, and Chris Volinsky. 2009. Matrix Factorization Techniques for Recommender Systems. *Computer* 42, 8 (2009), 30–37.
- [19] Amie Kreppel. 1999. What affects the European Parliament’s legislative influence? An analysis of the success of EP amendments. *JCMS: Journal of Common Market Studies* 37, 3 (1999), 521–537.
- [20] Amie Kreppel. 2002. Moving beyond procedure: an empirical analysis of European Parliament legislative influence. *Comparative Political Studies* 35, 7 (2002), 784–813.
- [21] Victor Kristof, Matthias Grossglauser, and Patrick Thiran. 2020. War of Words: The Competitive Dynamics of Legislative Processes. In *Proceedings of The Web Conference 2020* (Taipei, Taiwan) (*WWW '20*). 2803–2809.
- [22] Zoe Lefkofridi and Alexia Katsanidou. 2014. Multilevel representation in the European Parliament. *European Union Politics* 15, 1 (2014), 108–131.
- [23] Laurens van der Maaten and Geoffrey Hinton. 2008. Visualizing data using t-SNE. *Journal of machine learning research* 9, Nov (2008), 2579–2605.
- [24] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S Corrado, and Jeff Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems 26*. Lake Tahoe, Nevada, USA.
- [25] Monika Mühlböck. 2012. National versus European: Party control over members of the European Parliament. *West European Politics* 35, 3 (2012), 607–631.
- [26] European Parliament. 2021. Rules of Procedure of the European Parliament - Rule 180. https://www.europarl.europa.eu/doceo/document/RULES-9-2019-07-02-RULE-180_EN.html Accessed: 2021-02-14.
- [27] Georg Rasch. 1960. *Probabilistic Models for Some Intelligence and Attainment Tests*. Danmarks Pædagogiske Institut.
- [28] Soumya Sarkar, Bhanu Prakash Reddy, Sandipan Sikdar, and Animesh Mukherjee. 2019. StRE: Self Attentive Edit Quality Prediction in Wikipedia. In *Proceedings of ACL 2019*. 3962–3972.
- [29] George Tsebelis, Christian B Jensen, Anastassios Kalandrakis, and Amie Kreppel. 2001. Legislative procedures in the European Union: An empirical analysis. *British Journal of Political Science* 31, 4 (2001), 573–599.
- [30] European Union. 2021. European Data Portal. <https://www.europeandataportal.eu/en> Accessed: 2021-02-14.
- [31] Ali Batuhan Yardim, Victor Kristof, Lucas Maystre, and Matthias Grossglauser. 2018. Can Who-Edits-What Predict Edit Survival?. In *Proceedings of KDD'18*. London, United Kingdom.
- [32] Pengcheng Yin, Graham Neubig, Miltiadis Allamanis, Marc Brockschmidt, and Alexander L Gaunt. 2018. Learning to Represent Edits. In *International Conference on Learning Representations*.
- [33] Xiang Zhang, Junbo Zhao, and Yann LeCun. 2015. Character-level convolutional networks for text classification. In *Advances in neural information processing systems*. 649–657.

A TERMS PREDICTIVE OF ACCEPTANCE

A.1 Words Added

berec | fishing | should | equipment | office | registered | advisory | inserted | important | actions | 2018 | bargaining | best | therefore | transparency | regulators | fisheries | positive | withdrawal | plan | x | gender | financial | ppe | lisa | communication | defence | ” | fuel | second | external | toll | processes | common | buyer | skills | inform | reduce | digital | impact | 2005 | pension | v | contributions | support | council | fitting | agricultural | investigation | processing

A.2 Words Removed

berec | safety | eurojust | breeding | surveillance | area | council | consumers | human | 2 | authorised | powers | bodies | hosting | animals | articles | conditions | derogation | ; | 29 | medium | if | manufacturer | origin | allocated | | audit | implementation | provision | conformity | added | specific | 10 | plant | fitting | representative | action | fisheries | amending | 8 | current | financing | political | chapter | identifying | virtual | during | harm | compensation | breed

A.3 Context Words

” | appliance | appliances | controls | rco | prima | alternative | harmonised | threats | voice | egf | safety | eurojust | iccat | manufacturer | 63 | breeding | published | engines | customs | instrument | outside | associated | instructions | creditors | fittings | processed | destination | audit | uniform | number | cash | operating | notified | recipients | positive | institutions | appeal | alcohol | observations | berec | multi | _____ | guidelines | accounts | practical | expenditure | firms | ecosystems | saving

A.4 Title Words

community | customs | DDDD-DDDD | control | mediterranean | service | ’customs | supervision | installations | parliament | recovery | cableway | pollutants | multiannual | field | annex | equipment | temporary | council | documents | competition | anti-fraud | statistics | area | drinks | animals | appliances | DDDD/DDD | burning | gaseous | ukraine | fuels | it | policy | agency | zootechnical | plan | laws | ensure | other | fisheries | genealogical | spirit | financial | authorities | DD/DDDD | office | insolvency | investigations | management

A.5 Bigrams Added

their sector | opposition , | of meeting | the berec | , humification | this regulation | avoid social | this expert | way behind | a . | berec office | transmission of | in easy | were neither | . 2 | applicable the | violence is | eu ’s greenhouse | one of | risk premia | where applicable | economic operators | positive impact | within the | multinationals | at | relative deviation | institution , | is inserted | by sub | not properly | accept , | regions , | further amended | the third | complaint was | or federal | people and | , raising | family associations | intelligent mobility | - carrier | carrier economic | acoustic signals | , 51 | elected a | inserted : | signal processing | board of | 2018 prices | transport agency

A.6 Bigrams Removed

2 . | . 2 | . . | international efforts | , member | . where | the following | 5 . | hosting service | the case | ; the | human rights | the member | market surveillance | requirements of | data protection | subject to | in such | to that | 4 . | : the | the hosting | to be | which are | plan | ; evaluations ; | whether the | of that | covered by | and related | sharing and | the council | relevant for | eurojust shall | conditions , | . 3 | service provider | the development | provided for | the implementing | in hormonal | notified as | the efsd | that the | article 11 | any other | 6 . | of an | deleted . | take a

A.7 Context Bigrams

. ’ | : ’ | . those | 2 , | . 3 | . 2 | notified body | . ” | under other | authority to | requirements of | annual work | their citizenship | ’ the | management board | economic operators | is in | renewable energy | public sector | year . | 3 . | supervisory authorities | investment firms | quantified , | be deferred | the egf | and shall | regulation . | voice communications | this paragraph | promoter . | within a | , storage | / 22 | in other | shall : | of participants | authorities should | monitor the | of new | states remain | as set | ’ interests | financing types | year | , article 38 | commission in | consumers ’ | resident or | the institutions

A.8 Title Bigrams

council on | to regulation | cableway installations | supervision of | programme | multiannual recovery | 'customs' | of customs | , (| european parliament | the 'customs | recovery plan | and of | annex a | field of | parliament and | the reform | on insolvency | insolvency proceedings | a to | replacing annex | the field | and establishing | for cooperation | general budget | control equipment | customs control | budget of | rules applicable | and administrative | gaseous fuels | the mediterranean | DDDD/DDD on | financial rules | the use | in and | for trade | burning gaseous | procedures for | regulation of | the council | and supervision | to the | plan for | appliances burning | no DD/DDDD | zootechnical and | spirit drinks | medicines agency | european medicines

B TERMS PREDICTIVE OF REJECTION

B.1 Words Added

cabotage | these | deleted | ; | eu | societal | must | mercury | payment | illegal | benchmark | territorial | e | hydrogen | except | asylum | - | commercial | service | according | operational | include | basic | agreements |) | additionality | determined | case | consent | circumstances | after | ten | days | constant | negative | firearms | above | s | if | professionals | children | surveillance | set | only | settlement | amended | medical | hours | defined

B.2 Words Removed

energy | should | migration | additional | competitiveness | public | workers | corps | irregular | % | different | product | systems | international | forest | eib | efsi | remuneration | growth | joint | before | research | worker | economic | electronic | therefore | passenger | matter | works | solidarity | months | value | reporting | through | provide | online | can | eurodac | impact | monitoring | allowances | every | identity | account | cultural | concerted | supply | projects | format | structural

B.3 Context Words

_____ | allocation | rightholders | firearms | hosting | posting | allocations | resettlement | reserve | allowances | benchmark | foreign | core | free | labels | pnr | educational | verification | driver | collective |) | works | investments | 2030 | preservation | terrorist | forest | solidarity | remote | advanced | mercury | 25 | broadcast | ancillary | fingerprints | million | employees | condition | redress | settings | excellence | parental | 5% | penalties | travel | seller | exception | enisa | renovation | containing

B.4 Title Words

and | directive | market | DDDD | framework | services | DDDD/DDD/ec | requirements | </s> | protection | agricultural | energy | as | a | for | gas | decision | greenhouse | contracts | digital | operation | online | name | strategic | development | emission | regulation | view | of | from | at | georgia | instruments | establishment | structural | trading | record | plans | in | investments | supplementary | regards | the | specific | private | relevant | copyright | DDD/DDDD | pnr | posting

B.5 Bigrams Added

"10a . | communication , | normalisation process | welfare regulations | that activity | is deleted | become apparent | general production | ; | certificates were | and in | may propose | with the | and logistics | ' s | separation of | engine replacement | different generators | . in | hatred . | according to | or morality | as authors | made explicitly | value cases | , point | annex , | as jointly | lifting a | climate transition | - contributions | parties to | 32a is | valued by | place of | fuels for | directly awarding | a minimum | judges each | procedures overcoming | european union | ii may | new genetic | is amended | healthcare professionals | service + | , interpreters | leave may | between solid | status under

B.6 Bigrams Removed

no reason | and now | terrorism - | as the | contribute to | digital content | , possessed | other subject | guidance . | under the | electronic monitoring | the digital | . this | agricultural guarantee | than ten | legal body | which the | and other | , shall | 0 . | remote electronic | the cir | . member | the supplier | the passenger | authorised periods | solidarity corps | . in | in case | the product | ' association | of directive | policy objectives | the forest | - sharing | discussion . | the amount |] and | they are | intention or | 1 . | least likely | information on | same shall | - matter | state to | and of | carbon impact | union law | ; and

B.7 Context Bigrams

contents of | a sub | therefore , | hosting service | . _____ | the funds | of directive | pnr data | 000 for | parental leave | report within | regional operational | commission may | article 4 | investment board | paragraph 3 | produced from | procedure , | scientific evaluation | works or | their common | the driver | main third | or other | free allocation | state which | 27 . | deemed to | ' shall | states introduce | have given | . member | states may | data for | their rights | they shall | finance may | for free | programme's research | commission should | article 2 | - use | may request | down rules | consumers , | authority referred | the supply | which establishes | - and | be deemed

B.8 Title Bigrams

) and | and regulation | directive DDDD/DDD/ec | greenhouse gas | corps programme | rules for | european union | services in | , regulation | the eu | passenger name | eu pnr | by member | DDDD/DDD and | the structural | structural reform | name record | of passenger | data (| the framework | efficiency labelling | strategic plans | to georgia | for screening | european agricultural | DDDD/DDDD with | as regards | DDDD/DDD/ec and | record data | council amending | and weekly | DDD/DDDD as | daily and | directive DDDD/DD/eu | , and | a framework | pnr) | internal market | rules on | concerning the | screening of | protection certificate | for medicinal | supplementary protection | in criminal | , laying | against dumped | DDDD/DDDD on | of energy | of a