Both methods, meme and modeling, share a similar parallel to meme idea, theres an inheritance in citation of a potential meme and my text features for logit are the shared vocabulary, modeling tells me which is the inherited vocabulary.

Write something about how legal language invites text analysis methods because of how formulaic and standardized wording can get. Ambiguity is hashed out and words are weighed carefully and used in a relatively consistent manner.

In order for legal practitioners and scholars to have a correct understanding and application of existing case-law, their mapping of existing precedent ought to be comprehensive. Interdisciplinary endeavors towards creating such a mapping have involved the study of citation networks of court judgments from the U.S. Supreme Court, the World Trade Organization and the European Court of Human Rights, among others. Such quantitative approaches can reveal at once the motivations behind the citation of previous judgments and potential areas of the jurisprudence that have gone understudied in legal textbooks that attempt to catalogue important case-law.

Existing work on citation networks, either legal or academic, have focused on the possibilities of performing statistical modeling on the network. A dataset of connected texts, in the hands of a computer scientist, invites link prediction experiments. Can we predict which texts ought to be cited by future legal cases and academic articles? Or given a social graph, can we predict which users ought to connect? Literature on this task is plentiful and useful; lawyers or applicants to a court can use a recommender system to inform their search of existing case-law; researchers can contextualize their work or discover previous related articles.

Modeling a citation network can also divulge valuable information about the data, beyond what links ought to exist. The approach I take involves modeling the citation network of judgments from the European Court of Human Rights in order to reveal which textual features of the associated judgment texts are highly predictive of citation. The intuition behind this method relies on an assumption in which ngrams that are stronger predictors of citation are also descriptive of the court’s jurisprudence. A more extensive explanation of how this method and how it is evaluated is fleshed out below.

One distinction from this comparison, with important methodological consequences, is the lack of an official obligation on judges of the ECHR to be formally bound by any existing precedent (which one often finds in other constitutional and supreme courts). A reference to a previous judgment by a judge is not necessarily central to the legal argument in the judge’s opinion. The impact this may have on a model of a citation network and its associated judgment texts is not negligible. My experiments hope to reveal patterns in the shared vocabulary of judgments linked in a citation network, which could then be associated to literature on that area of the case law. It is possible that jurisprudence develops without any detectable thread along past cases, as the facts at hand officially have more weight in the judge’s decision-making. Previous research on the court does however show judicial practices motivated by internal consistency, with references to existing case-law aligning with what one might call “de facto stare decisis” (precedent is respected despite no obligation to do so).

Background

History of the European Court of Human Rights

Officially inaugurated in 1959, the establishment of the European Court of Human Rights is best understood as both an institution within a nascent legal system and a political project born of the interests of Western Europe and liberal democracy. The court is meant to apply the articles of the European Convention for the Protection of Human Rights and Fundamental Freedoms, entered into force in 1953. Initially, the court operated in a system that was legally underdeveloped (human rights in Europe) and strongly influenced by national interests. As the court’s autonomy grew, a more complete legal system protecting human rights within Europe eclipsed foundational concerns for diplomacy and the threats of fascism, Soviet imperialism and the extensive human rights violations witnessed during World War II. Today, the combined instruments of the convention and the court, understood as a regime, is considered the most effective in the world CITE. The legal dynamics of the regime draws comparisons to a bill of rights (the convention) interpreted and enforced by a nation’s constitutional court (Strasbourg) CITE2. Any of the 800 million individuals from the 47 countries that are party to the convention can appeal to the court when their rights have been violated by their government, hoping to be compensated financially (what the court calls “just satisfaction”) or to bring about a change in national laws that contravene the convention. CITE THE EUROPEAN EXPERIENCE The court’s jurisdiction has grown extensively since the convention’s ratification in 1953 by a mere 10 states. The fall of the Berlin Wall has seen the membership double, with former Soviet bloc countries become signatories. This history provides for a special development of how Europe came to develop a system to protect human rights, and it is partially outlined below.

FOOTNOTE this One distinction from this comparison can be found in later pronouncements by the court establishing the so-called fourth instance doctrine, namely that the court should not act as final court of appeal, reviewing national courts’ potential errors of fact or law.

CITE:(A Europe of Rights: The Impact of the ECHR on National Legal Systems).

CITE2 **L**aw of the European Convention on Human Rights

A history of strategic Cold War concerns defined the scope of rights considered in the convention. A set of human rights covering political and civil life are typically referred to as the first generation of human rights, finding their genesis in political thought of the 17th and 18th century. The majority of the rights protected by the convention fall under this class, such as the right to property, freedom of expression, liberty, life and privacy. The omission of rights belonging to the second generation of human rights in the drafting of the convention illustrates some of political priorities of the signatories. In response to accusations leveled at governments of the Soviet bloc, a definition of human rights including social and economic rights such as free healthcare, housing, education and a right to work were promoted by communist regimes and their allies. The focus on the first generation of rights reveals the birth of a human rights regime as a product of contemporary political forces:

“…human rights had become a direct battle line in the cultural Cold War of the 1950s and 1960s, seeing the opposing camps of, respectively, the CIA-funded International Commission of Jurists and the Moscow-oriented International Association of Democratic Lawyers fiercely battling over the concept”

Early pragmatic concerns of legitimacy and precautions taken to not appear as a threat to national interests also advanced certain concepts within the case law. The development of the doctrine of ‘margin of appreciation’ is arguably a result of these priorities. A central principle within the jurisprudence of the court, margin of appreciation is a legal device employed to reconcile the impossibility of a universal application of convention articles given specific circumstances a member state may be facing, such as a national emergency or a unique moral landscape. The doctrine allows for a certain amount of discretion on behalf of the member states if justified by the particular conditions at the time of the purported violation.

With its legitimacy gradually consolidated, ECHR judges were better positioned to advance jurisprudence in bold directions that allowed for more progressive interpretations of the convention as well as stronger expectations from member states’ legal systems. The interpretation of provisions became more forward-looking with a principle establishing that the ‘purpose and object’ of the convention be given primacy when the spirit of a case fails to be captured by the exact wording of any provision. The convention was subsequently defined as a ‘living instrument’ by the court, allowing for a more dynamic interpretation. Member states were also expected to provide individuals not only with an abstract recognition of the rights conferred by the convention, but what would be described by the court as a ‘practical and effective protection’ of these rights.

WHAT RIGHTS

PRECEDENT/CASE LAW

WHAT A JUDGMENT LOOKS LIKE

Related work

SOME TITLE THAT SUMMARIZES THIS PAPER

The article ‘Empirical Studies of the Webs of International Case Law: A new Research Agenda’ maps out a variety of methodologies employed by political scientists and legal scholars to examine the ***developme***nt of case law, especially in international courts, both qualitative and quantitative. The authors propose an approach that combines citation network analysis, computational linguistics and legal analysis as a new direction for international case law scholarship.

The article sets two branches of case-law scholarship in opposition. One branch represents the political scientist’s approach, which operates empirically and quantitatively to focus on the behavior of judges or the outcome of cases, often ignoring the content of the judgment itself. What’s in focus is perhaps the judge’s professional or ideological background, or the court’s motivation for the number of citations used. The other branch is represented by the legal scholar. The latter can be situated in the ‘text book tradition’, updating a record of the important cases pertaining to an area of the law. A legal scholar may also study an important decision in order to clarify the precedent that has been set. The authors point critically to the manner in which many legal scholars are too eager to find congruence between previous practice and legal categories, and what might arise of new case law.

A brief explanation of citation network analysis, its pertinence to the study of the law, and various importance measures is followed by an overview of how citation network analysis has been applied to domestic and international court jurisprudence. A legal theoretical framework is set out in which the interpretation of the law by practitioners and scholars requires a mapping of the case-law that must be comprehensive in order to be correct. The tools outlined in this article provide support for a theoretical framework highlighting the importance of case-law over policy in the application and interpretation of the law.

Despite acknowledging the weaknesses of network analysis (a static picture of the law is posited, and the reason for citation is ignored) the authors advance an approach that builds on previous attempts to combine quantitative methods such as network analysis with qualitative examinations of legal language and argument. More specifically, they perform modularity maximization on a citation network of the Court of Justice of the European Union. This technique, which groups cases into communities according to shared connections, provides the authors with a finding that a central node to one community, ranking high in authority, has gone largely ignored in textbooks. Having made note of the oversight detected through network analysis, an analysis of the case’s language is anchored on the presence or absence of a passage involving the word ‘effectiveness’. They employ corpus linguistics tools to show that different collocations involving ‘effectiveness’ can be juxtaposed with case-law network communities, presenting a systematic study of precedent combining network structure and language.

SOME TITLE THAT SUMMARIZES THIS PAPER

The article ‘Who Should I cite? Learning literature Search models from Citation behavior’ is heavily used as inspiration for my thesis and comes from the field of Natural Language Processing and Information Retrieval. The authors combine information from a citation network, the text content and metadata to produce a retrieval model for the sake of scientific literature search. Their features include:

-A score for the similarity of terms in the query (an article abstract) and the documents (article reference canditates), relying on tf-idf values of the vocabulary.

-The PageRank score of a document, and its citation count, as well as the citation count of its venue, its author, and its author’s H-index.

-The number of years between the article abstract (the query) and the potential reference.

-A similarity score for query terms and document + citing documents terms.

-PMI, EXPLAIN IF I USE IT

-A cosine similarity score for topic distributions vectors of query and document , and a score for query and citing documents

-A series of features based on the query author’s previous citations, and the names of the author(s) of the query and documents.

They compare different classifiers for learning the weights of these features, using an iterative approach where sample articles are added to the training data at every iteration, and being compared the training abstract’s actual reference list. The classifier iteratively learns what is relevant and what is not and updates the feature weights accordingly, before resampling and retraining. And SVM classifier outperformed a L-BFGS classifier evaluated with mean average precision. Using all features also produced the best results, with citation-count, and topic similarity between abstract and reference candidates(+ reference’s citing documents) AND TERMS PMI??? Proved most important.

SOME TITLE THAT SUMMARIZES THIS PAPER

The article ‘The Role of Precedent at the European Court of Human Rights: A Network Analysis of Case Citations’ examines the development of precedent at the ECHR and the various political, strategic and legal motivations that factor into the courts decisions.

Beyond ensuring a consistency in judgments, a court can use precedent in order to legitimize its decision-making and convince parties to follow its judgments and principles. In the case of an international court such as the ECHR, examined by the authors, affected parties are state actors. The court has no mechanism for enforcing its judgments so a domestic court needs to be persuaded by resorting to the legitimacy of existing precedent. A third type of citation practice is said to be influenced by a careful consideration of a nation’s interest and how differing morals play against individual rights. This caution is also strategic; the display of sensitivity is necessary to influence the affected government. These three perspectives (consistency, as well as ‘strategic’ and ‘relativist’ legitimation) present varying expectations as to why judges cite precedent, how important the cited cases are, and how many cases are cited. Hypotheses as to use of precedent are formulated to test the validity of either perspective.

The article goes on to describe the data (an ECHR citation network ending at 2006), the distribution of inward and outward citations and their totals over time, how this compares to the US Supreme Court and other figures, i.e. the number of citations and cases, and the distribution of cases across convention articles.

The authors arrive at a series of findings that support a perspective in which citation is done to communicate with the involved parties. Hypotheses based on a view in which the purpose of citation is to persuade domestic courts and governments are validated by their research. By examining the characteristics of the cases in the network along with number and importance of cases they cite (hub score), they find that 1) violations of convention articles associated to more fundamental issues such as life and torture are backed up by more authoritative cases, 2) these judgments also go on to be more authoritative, i.e. cited heavily 3) the court relies on precedent more when ruling against a government.

Modularity maximization is performed on the citation network to create communities of citations. They find that the communities are formed around common legal issues rather than being grouped by nation, confirming that 4) the court cites previous cases based on legal substance rather than being concerned with legal culture of a domestic court system and other country-specific characteristics.

SOME TITLE THAT SUMMARIZES THIS PAPER

The article ‘Predicting a Scientific Community’s Response to an Article’ predicts downloads and citations of articles based on their text content rather than citation network structure as past approaches have. They employ supervised learning methods to model the impact of articles, with different modeling approaches for either forecasting or predicting within the same time span as the training data. The prediction task also differs for two datasets; linear regression is used for economics articles whose total number of downloads is modeled whereas a citation network of computational linguistic papers is modeled with logistic regression (either cited or not, three years after publication).

The authors go over generalized linear models and penalized regression, especially l2 penalty (ridge regression), which is employed for the non-forecasting predictions task. A novel time-series regularization is applied for the forecasting of citations and downloads. Having split their data by year, the subsequent coefficients for every feature and every year are regularized to control for overfitting as well as variation across years.

EXTRAPOLATION?

FORECASTING RESULTS , features

The advantage of the time-series regularizer is the interpretability of coefficients that captures the importance of ngrams or metadata to the model (and therefore to the scientific community) and its fluctuation over time. The regularization penalty on large variation from year to year results in a smoother timeline, corresponding to research trends, than that of multiple models trained on year-separated datasets (and more informative than the steady coefficient of a model trained on all the data at once). The authors also compare the timelines of a handful of coefficients to timelines illustrating the frequency of these terms in the same literature. Their analysis favors the trends produced by their method, as terms like ‘parsing’ and ‘grammar’ show more realistic developments.

SOME TITLE

The article ‘Inheritance Patterns in Citation Networks Reveal Scientific Memes’ employs the understanding of a meme as developed by biologist Richard Dawkins; gene replication and propagation through the gene pool is analogous to a keyword or phrase’s distribution in a citation network (as one might also see the sharing of music, ideas or images in culture). Documents cite each other creating an academic ‘bloodline’, in which certain terms may be ‘inherited’. A meme refers to terms that are inherited in a more consistent manner, much like genes.

The ‘meme score’ described by the authors and calculated over various scientific article citation networks is an attempt to identify ngrams that are at once important to the literature while also being interesting. Importance is merely how frequent the meme is, and it is considered more interesting if it has a higher ‘propagation score’ in the network. A more detailed explanation of this measure is included in the methodology section below.

The authors provide a summary evaluation of this measure through an examination of both frequency ratios and propagation scores of all potential memes across datasets. The similarity in distributions of memes (highly clustered around high propagation and mid to low frequency) is lost when a network is randomized, with nodes assigned random texts. Meme distribution therefore illustrates a process which time and network structure cannot explain on their own. A handful of terms independently hypothesized by the authors as memes, such as ‘quantum’ and ‘fission’ all share a combined high frequency ratio and propagation score. A more in-depth evaluation follows.

The top 50 high-scoring physics memes are mostly deemed relevant to the field by expert annotators, and many appear in Wikipedia physics articles. In a high level of agreement, the two annotators judge the 150 memes with the highest meme-scores to be relevant at a much higher rate than meme’s selected at random.

A list of ground-truth memes extracted from Wikipedia article titles and links on physics and science is compared to the list of the 5178 highest scoring memes, as well as lists produced with alternatives to the meme score. The alternative metrics are applied in order to examine whether alignment between Wikipedia ‘memes’ and results of their meme formula can be reproduced with measures based on different criteria. Scores based on frequency alone, frequency changes over time, and frequency difference across journals exhibit much lower agreement with Wikipedia ground-truth memes. Wikipedia memes also tend to have very high meme score. The authors end with a call for meme identification using more advanced language processing methods, in order to track the inheritance of memes beyond ngram matches.

DATA:

The entirety of the data used for my experiments consist of a corpus of judgments texts from the European Court of Human Rights and a network representing these judgments and the citations between them. These two datasets were generously provided by the iCourts research centre at the University of Copenhagen. Additional datasets employed in my analyses are a corpus of publications covering the case law related to particular articles of the European Convention of Human Rights, as well as a list corresponding to ‘controversial’ issues from Wikipedia. ADD IF ADDED

The judgment texts total \_\_\_ documents, mostly in HTML format. A look at how these judgments are distributed across time reveals a much more active court in recent years: \_\_\_ percent of the data consists of judgments from after 2010, \_\_\_ percent from after 2000 while the first 20 years of the court’s history produced \_\_\_\_ judgments, representing \_\_\_\_ percent of the corpus.

An important preprocessing step involved the removal of texts corresponding to the court’s decisions rather than judgments. The former purely examines the admissibility of the case and will consequently not go at length into the merits of the case. Every judgment text is made of several parts, each varying in degrees of standardized language. The opening of a judgment lays out the respondent, applicant and which judges are sitting for this particular case. A ‘Procedure’ section establishes when the application originated and the article of the convention that may have been violated. Under the ‘Facts’ heading, the circumstances of the case are explained, often including the trajectory of the case along various national courts. Under the same header, the court will sometimes look at relevant law from the respondent state as well as international law beyond the European Convention on Human Rights. A section entitled ‘The Law’ examines the purported violation of the applicant’s right, (as it is argued by the applicant and the respondent Government) how the article convention applies to case at hand and how the court has dealt with similar cases in the past. If the case is submitted to the court under many article violations, each violation and article is considered in turn. The court also considers the extent and nature of the compensation awarded to the applicant if a violation of a convention has been found. Any judge(s) disagreeing with majority opinion of the court will lay out their argument in a dissenting opinion, and possibly in a separate opinion, if a judge is dissenting for reasons that differ from other dissenting judges. Judges can also express their view in ‘partly concurring and partly dissenting’ opinions.

Although it is possible that judgments may cite previous cases due to these dissenting opinions, I have filtered out these sections from the texts when modeling the court’s citations. The sections of a judgment that approach legal boilerplate were kept in with the assumption that a proper model would ignore such linguistic patterns on account of their pervasiveness.

Having removed from this corpus the dissenting opinions, decisions, documents in other languages or formats other than HTML, the data is converted to plain text and any reference to a case number or case title is also filtered out so as to avoid predictors that simply result from the citing and cited documents having in their shared vocabulary the title or number of the cited case. In some instances, a case number may correspond to multiple documents, if the court revisits the case in order to handle compensation for a violation, which has not been dealt with. For the sake of my experiments, these resubmissions are ignored. When modeling the texts and their citations I am also limiting my data to judgments that appear in both the citation network and the corpus of judgment texts. After these preprocessing steps, the total count of judgments used is \_\_\_\_\_\_.

The citations between judgments are represented by a directed graph, upon which the modeling of citation is performed. This network has been constructed at iCourts by identifying a variety of citation patterns in the texts: a case number or title is referenced, or paragraphs from the cited text might be mentioned, often in the form of “(see paragraph…)”. The graph is composed of \_\_\_\_ nodes and \_\_\_\_ directed edges, where source nodes are the citing case and targets the cited. Node attributes consist of various metadata, some of which is used as features for prediction. The language of the judgment (English or French), the type of document (judgment or grand chamber judgment) and a short string identifying the conclusion reached by the court are all ignored in my feature candidates, as are title and originating body. My experiments do make use of the case’s date, respondent country and the article(s) under which the application is submitted.

The distribution of cases across countries and across articles is far from uniform. A total of \_\_\_\_\_ cases (\_\_\_ percent of the data) fall under article 6, protecting the right to a fair trial. Italy is the respondent country in \_\_\_ percent of the cases, \_\_\_ of which deal with article 6. Russia and Turkey are also overrepresented (\_\_\_ percent and \_\_\_percent). The court has classified cases according to their level of importance (1,2, 3 or 4), corresponding to the case’s contribution to the case-law. The relative frequency of certain respondent countries or purported article violations can be indicative of applications to the court where the circumstances of the case closely resemble those of existing case-law, and consequently provide little opportunity for the development of the case-law. Their estimated importance is thus also used in this project’s experiments.

I have created article–specific subgraphs for articles that appear at least \_\_\_ times, both as a result of the observing this distribution across articles, and in order to return ranked predictors that would hopefully be more specific to the case-law for each article, while taking into consideration the sufficiency of case numbers and graph edges for adequate modeling. The majority of the cases fall under more than one article, and an analysis of the most common article pairings may be of interest, but I have limited my experiments to articles:\_\_\_\_\_\_. The number of edges, representing citations, in these subgraphs range from \_\_\_ to \_\_\_\_\_.

I have collected a corpus of relevant publications that are affiliated to the court without being binding on its judges. One series, Guides on the Convention, is prepared by the Research and Library Division within the Directorate of the Jurisconsult and is made available on the ECHR website. A second series, Human Rights Handbooks, is made available by the Directorate General Human Rights Council of Europe. Both series cover important cases, discuss how the convention article has been interpreted, and provide insights on how the article ought to be implemented. I have collected \_\_\_ of these PDF documents, which I have converted to plain text files. The average publication contains \_\_\_ pages, \_\_\_ words.

Article-specific publications from the series often contain in the title ‘A guide to the implementation of Article \_\_\_’.

Wikipedia provides a list of topics that it deems controversial on account of the number of revisions Wikipedia contributors carry out and how edits often are performed in a circular manner: a piece of information is changed only to eventually be re-edited back in, in a circular manner. This list of controversial issues is divided into \_\_\_ sections (such as Philosophy, Religion and Sexuality) and at the time of writing totals \_\_\_\_\_ issues.

Methodology

I use stability selection combined with L1 logistic regression models to identify variables that are highly indicative of citation. Stability selection is essentially a randomization technique involving the application of a selection algorithm, such as logistic regression. Rather than being applied to the whole dataset, stability selection re-estimates the model many times over different random subsets of the data, drawn without replacement. After many iterations, the results are aggregated by detecting which variables are selected at a higher frequency over the subsamples.

Logistic regression is a discriminative model that estimates P(y|x) where output is a binary variable. The inner product of the feature vector f and weights w is the first step towards estimating P(y|x).

Sum wifi

This maps f to to a real number, between -∞ and∞. Including this inner product in a logistic function allows for the mapping to fall within the range of [0,1], resulting in actual probabilities.

p(y = 1|x) = 1/ 1 + exp(−θ>x) .

The estimation of unknown parameters in a logistic regression model is done with conditional maximum likelihood estimation. The method chooses values for the regression coefficients (w) that maximize the probability of the training data.

max θ nSUM i=1 log p(yi |xi , θ).

The maximization will overfit the training data; our model will pick up on noise and find non-existent patterns. Regularization keeps overly large weights from being learned, by adding a penalty to the maximum likelihood estimation equation, penalizing complexity and controlling the variance. The penalty term here is R(w)

wˆ = argmax w X j logP(y (j) |x (j) )−αR(w)

The regularization term R(w) in Ridge regression is the l2 norm of the weights. A higher penalty term leads to lower weights, but never allowing any of the coefficients to get to zero. Thus variables selection is not an option with Ridge regression

*wˆ* =argmax􏰁logP(y(j)|x(j))−α􏰁w2i

L1 regularization, or lasso (Least Absolute Shrinkage and Selection Operator) regression involves a penalty term that sums over the absolute values of the coefficients rather than summing over the squares of the coefficients. In this case weights can shrink to zero, allowing for variable selection, although optimization is more difficult.

SOME EQUATION

L1 provides for sparse weight vectors, many of them set to zero. Fewer features will be used those that correlate are competing for selection. Only one feature will be chosen and one will be ignored. This is overcome with stability selection, as the model is reestimated many times, resulting in a set of stable variables­

Features

The estimator upon which stability selection is performed in order to return ranked predictors is limited to training on features whose interpretation can help explain the court’s case-law. Thus, features pertaining to textual similarity between judgment pairs, citation network metadata and node importance measures are left out. However, in attempting to gauge the performance of a predictor trained on binary ngram feature vectors, judgment citations are modeled by including these additional features.

One such feature involves the textual similarity between citing and cited judgment, based on the cosine similarity of either document’s tf-idf weighted vector representation.

**Tf-idf term weighting**

Tf-idf term weighting is a method for producing vector representations of documents whereby the terms that are important to that document are given a stronger weight at the index corresponding to that term (instead of a 1 for presence and a 0 for absence). ‘Tf’ refers to the frequency of the term in the document, which will be multiplied by the ‘idf’, the inverse of that term’s frequency in the entire corpus. This controls for words whose high frequency in a document should be unsurprising: common words such as determiners or words such as ‘court’ and ‘applicant’, which should be found in every document in the corpus. The weight should be higher if a term appears frequently in a document and if it is specific to it.

EQUATION LOG?

A document is represented as a tf-idf vector, with a length corresponding to the total number of tokens in the corpus, and zeroes where a specific token is absent from the document. The calculated tf-idf value is stored at the index for every term present in the document.

The similarity between the vectors for a cited and citing pair is measured by calculating the cosine of the angle between either vector.

EQUATION

The numerator of this function is the dot product of either vector, and the denominator is the product of the vectors’ Euclidean norms. The cosine similarity measure will be 0 if no words between both documents overlap and 1 if two documents are identical.

**Network analytics**

The authority score of the cited judgment is included as a feature as well. This is returned by running a Hyperlink-Induced Topic Search (HITS) algorithm over the citation network, producing both authority scores and hub scores. A node with a high authority score is cited by more nodes with higher hub scores. A node with a higher hub score cites more cases with high authority scores. The eigenvector centrality score of the cited judgment is also included. The eigenvector centrality of a node is derived by the eigenvector centrality of its neighboring nodes. A case has a higher eigenvector centrality score than one with the same amount of MAKES SURE I GOT The RIGHT SCORES. The indegree centrality score included is simply a count of the node’s incoming edges.

METADATA:

Four more features included in the full feature set make use of metadata associated to the nodes of the citation network. I’ve included a feature for the number of days in between the cited and citing judgment, assuming that predicting a link between two judgments ought to be informed by whether the judgment have happened too close together in time. Whether or not the respondent country in the case being cited is the same as the citing judgment was included as well, under the assumption that case law pertaining to the same country might be highly relevant to future judgments. While my modeling is performed over article-specific subgraphs and their associated documents, I am including as an additional feature the number of articles under which the judgment falls that both cases have in common. A case about the right to privacy as well as protection from discrimination should be more likely to cite cases that are submitted under the same articles. Finally the court provides a rating for almost every case, from one to four, indicating its importance to the case-law. The court’s importance rating for the cited case in a pair of judgments is used as a feature as well.

Inspired by CITE, an additional measure of an ngram’s importance within the case law is its ‘meme score’. In this instance, no modeling of the citation network or text is performed. The score is calculated by multiplying the ngram’s frequency in the subcorpus (associated to the case law covering an article of the convention) and a novel metric proposed by the authors called propagation score. It is essentially a ratio of the judgments that cite case law containing the ngram over judgments that do not. The propagation score is represented by EQUATION #

jm->m/j->m + δ/ jm->n +δ/j->n + δ

the numerator is a ratio of judgments j that contain the ngram (or meme m) and cite at least one judgment containing the meme (represented by jm->m) over all judgments that cite another containing the meme, j->m. The denominator is meme-containing judgments whose citations do no contain the meme (jm->n) over all judgments that do not cite judgments with the meme (j->n).

CITE has experimented with different levels of controlled noise δ to both perform smoothing as well as filtering out high scoring memes due only to their overall frequency. Their experiments show that δ = 3 provides best results for their data, and for expediency’s sake the same controlled noise will be used in my work. The final meme score is calculated by EQUATION #

Mm = fmPm

A given meme m is derived by multiplying its relative frequency in the subcorpus with its propagation score. The frequency ratio does not total all counts of the ngram but rather the number of judgments containing the ngram over the number of total judgments.

Evaluation survey?

Experimental set up

For every article of the convention and its associated subgraph of the citation network and judgment texts, stability selection is performed in order to retrieve a ranked list of terms (unigrams, bigrams or trigrams) corresponding to predictors of citation. In order to validate the scores attributed to text features, this ranking is compared to the appearance of predictors in the literature on the case law of the court. To get an idea of why citation happens, the ranking is also compared their appearance in either the ‘Fact’ section or ‘Law’ section of the judgments. ***And in a list of ‘controversial topics’***. The performance of the predictors trained on ngrams only is also compared to predictors trained on textual content, metadata and network structure features.

I am using a module for feature selection provided by SciKit-Learn named RandomizedLogisticRegression, which performs stability selection and returns a list of all feature scores, between 0 and 1. This is performed on features corresponding to a ‘Bag of Common Words’ between two judgment texts. Rather than vectorizing each judgment text from the article-specific subcorpus, every feature vector used for modeling describes word occurrences (1 for presence, 0 for absence) shared by two documents, where the length of the vector represents the length of the corpus’ vocabulary, with stopwords removed. The intersection of the vocabulary of all texts associated to all connected nodes in the article-specific subgraph makes up half of the data, while the second half is an equal number of vocabulary intersections from pairs of judgment texts sampled at random without replacement, whose nodes in the subgraph are not connected. The prediction task is thus a simple binary classification, where a judgment is predicted to cite another based on the intersection of their vocabulary. Performing feature selection on this prediction task to improve the accuracy of the estimator is what returns terms that are stronger predictors or citation. Whether these feature scores correspond to relevance and importance in the case-law is the purpose of the following experiment.

Ranked predictor lists are returned for text features as unigrams, bigrams and trigrams and the following validation step is conducted with every ngram representation.

WHAT SHOULD I DO

Judgment texts are then split by section: the court’s description of the circumstances surrounding the case (‘The Facts’) and it’s interpretation and overview of legal issues arising from the case (‘The Law’). The score attributed to every text feature is compared to the count of that ngram in either section of the judgment. Similarly, CONTROVERSIAL

In order to gauge the performance of the predictor used to return these feature scores, an assortment of features related to the data’s textual content, the citation network’s metadata and its structure is used for training and comparison. The \_\_\_\_\_\_ error metric !! is returned on all possible combinations of these features. The ngram representation of the judgment pairs vocabulary intersection is used as a baseline, onto which all other combinations of features are added.

Prediction accuracy, what kind of error measure. Different features

Evaluation survey?

“There are 193 trigrams (three word sequences) I would like you evaluate. I would like you to answer Yes if you feel that the expression might be important to Article 10 -Freedom of expression- case law. Answer No if you do not see any relevance to the case law, or if you feel the expression would be so pervasive in court judgments that it is not interesting enough.   
  
My process involved removing stop words, such as 'by', 'to', 'for', etc. This will require that you attempt to infer what may have been in between the words of the trigram, if anything. For example, a trigram extracted from P-1 case law is 'delay paying additional'. This is found in the judgments texts as   
"the national authorities' delay in paying the additional compensation” but other formulations are possible.”

Article 10: Freedom of Expression

* 1. Everyone has the right to freedom of expression. This right shall include freedom to hold opinions and to receive and impart information and ideas without interference by public authority and regardless of frontiers. This article shall not prevent States from requiring the licensing of broadcasting, television or cinema enterprises.
* 2. The exercise of these freedoms, since it carries with it duties and responsibilities, may be subject to such formalities, conditions, restrictions or penalties as are prescribed by law and are necessary in a democratic society, in the interests of national security, territorial integrity or public safety, for the prevention of disorder or crime, for the protection of health or morals, for the protection of the reputation or rights of others, for preventing the disclosure of information received in confidence, or for maintaining the authority and impartiality of the judiciary.

The second volume of the human rights handbooks, “Freedom of Expression: A guide to the implementation of Article 10 of the European Convention on Human Rights”, goes over the important case-law surrounding this article and its associated doctrines, principles and concepts. The intended audience of the guidebook is judges, both at national levels and at the ECHR, as the case law is summarized for the purpose of guaranteeing a level of consistency across member states and within the court itself. The guidebook begins with some general considerations on freedom of expression and goes on to evaluate the manner in which judges have interpreted both paragraphs of article 10. The following summarizes the guidebook, laying out the potential keywords and phrases that seem to reflect core concepts in Article 10 jurisprudence.

Freedom of expression aligns with **freedom of assembly**, while being potentially at odds with rights such as the freedom of **conscience and religion** and the right to **respect for private life.** Jurisprudence on freedom of expression has weighed conflicting rights with an emphasis on how freedom of expression is important to **democratic society**, **progress and individual self-fulfillment.** Also important is the **pre-eminent role** of the **press.** Expression is protected regardless of its content, unless what is expressed negatively affects other rights and freedoms, or invites racism, hatred or Nazism.The court is then faced with the **paradox of tolerance,** i.e., can intolerance become the result of allowing absolute tolerance in expression.

When a state has interfered with freedom of expression, considerations that may contribute to a justification of this interference include the **type of expression, the means of dissemination, the audience,** and the **truth** of the expression.

The freedom to hold opinions, established in paragraph 1 of Article 10, also leads to a freedom of **not being forced to express** this opinion. The paragraph 1 freedom to impart information and ideas involves concepts such as the **public interest, artistic freedom** and how **creative work** shapes **public opinion**. The state cannot interfere between the media and its audience.

The court also considers how **the truth of value judgments** differs from **the existence of facts** and how this applies to a **crime of insult** and **defamation**. It looks at the **proportionality of an interference**, and whether there is enough truth to the expression. The right to receive information involves the right to be **adequately informed. Freedom of the press** andits role as a **political watchdog** ties in with ideas of **political debate** and **political issues,** and these ideas are weighed when the press may affect the **reputation of others** through **rumours and allegations. Objectivity** and **the protection of sources** are other factors. Paragraph 1 covers the right of the State to require broadcasting licenses. This prevents **monopolies** and ensures a **a diversity of information**, responding to the needs of the public

The jurisprudence is particularly concerned with protecting expression that might **offend or use strong language, regardless of the form of the expression – whether in writing, spoken or in picture. The court has drawn the line at expression that incites violence.**

Freedom of expression is therefore not an absolute right, and the second paragraph of Article 10 covers potential justifications for restrictions on this freedom. Civil servant and state employees are not free in the same way as private citizens due to responsibilities and status stemming form their profession. **Judges** are expected to preserve their **impartiality** by **limiting themselves in their public expression**. **However, confidentiality** can only be expected of civil servants for **particular categories of information**. Information can only be restricted if national security is actually threatened. **Journalists** are other professionals who have responsibilities such as ensuring that the information they share is accurate

Paragraph 2 is concerned with the different ways in which a state may restrict freedom of expression, whether **conviction**, forcing to **reveal sources, fines** or **censorship**, which is a special concern as it restrains expression before it happens. Certain punishments have led to an examination of the government’s **dominant position** and the **proportionality of the interference** on freedom of expression. The Court has expressed a desire to **protect sources**, as they may avoid talking to the media if they lose their anonymity. Censorship as well as fines and trial expense may also impede on the work of press unjustly.

The second paragraph outlines criteria for interfering with freedom of expression, such as protecting the reputation of others and national security. The court has however emphasized a strict interpretation of these criteria, in that restrictions must be based on these criteria only. Any law restricting freedom of expression has to be published and accessible, and precise enough for individuals to govern themselves accordingly.

The image or honor of a country is not covered as a reason for interference, nor are state prestige and authority. The reputation of high officials should not be protected with higher penalties. And threats to national security must be real and not an uncertain possibility.

Interference as a means to restrict freedom of expression must have an aim that is proportional to the means of the interference. It must respond to ‘pressing social need’. The second paragraph lists potential aims, and jurisprudence has developed to define them.

When the aim is the protection of national security, the court has found that confidential information may be restricted, but has found that once the information is released and has been received by the public, further limitations on sharing its content is unjustified. The court also considered whether the interest of the public in having the information is more important than national security. On issues of conflict in Turkey, has differentiated writing that encourages violent means from communication that criticizes the state’s actions.

When a government acts in a way to prevent disorder and crime, the court examines similar considerations. Was the expression an incitement to violence or terrorism or an exercise in scrutiny of the government, and was the punishment disproportionate? Can military personnel be punished for engage in a discussion of ideas about the army, or does that threaten the system of defense.

As for restrictions based on protecting **the reputation of others**, the court has prioritized the **freedom of the press** and **political debate** over the reputation of politicians, who are expected to be criticized. **Pubic discussion** depends on **media criticism** and either **censure** or punishments after an article is **shared** may have a **chilling effect**. Nor can a journalist’s **value judgments** be tested on their **truth**. A potentially offensive article can be protected by concerns such as the “**free debate of ideas”** and exaggeration can be allowed**.** Whether or not a journalist intended to defame someone is another consideration as are the **purpose** and **impact** of an article. Civil servants, on the other hand, were considered not to carry the same expectation of being closely scrutinized.

When the court has weighed freedom of expression over the right to religion, it is expected that religion can be criticized and denied. But expression that offends the religion without any particular purpose is condemned, such as portrayals of religious objects in an unnecessarily negative offensive light.

The court has also found that restrictions on racist expressions would not be upheld when the purpose is not propagation, but to expose and analyze these views through journalism, with no say in how balanced or objective the report may be. The court has also protected expressions against judicial authorities, which may have been limited to preserve public trust. The court recognized the impossibility of preventing prior discussion of disputes, outside the judicial system, which it argues cannot operate in a vacuum. Criticism of judges’ independence or impartiality should not automatically be restricted either.

An additional survey of the jurisprudence is found in an older document, put out by the Council of Europe Publishing, covering the case law of Article 10 CITE. I have included below the approaches and principles expressed in judgments that this publication outlines, if it has been excluded form the human rights handbook summarized above.

It is noted that the freedom of expression cannot be restricted in varying manners, according to publication’s country of origin. A restriction on the a psychiatric patient’s right to reading material and entertainment was found to violate his rights as it was not based on any law of the member state. When a law demanded ‘good behavior’, it was found to be insufficiently precise to guide one’s actions.

A survey of the case law on media freedom reveals that this area has generated considerable jurisprudence. A restriction of expression is not necessarily a violation of that right if it is meant to punish the abuse of that freedom by military personnel writing in the media. The case law emphasized revolves around similar notions of freedom of the press, their role in forming the public’s opinion and how this allows criticism of the state and state actors. A politician must allow, for example, that journalists bring up their criminal pasts. The court had found no violation when a member state convicted a journalist for being overly and unethically critical of a judge. Punishment could not be excessive however; an award for damages of 1.5 million pounds was condemned by the court. If a publication influences criminal proceedings, the court has to consider the degree to which justice is impeded. The publication of photos of a suspect in a criminal trial cannot be prohibited, especially if it is trial of public concern. A professional’s career must not come before the interest of the public and their right to know about the professional’s potential incompetency.

The court had, as of the date of this document’s publication, been relatively unwilling to find a violation in cases of access to information. Applicants have appealed to the court for protection of their right to information on themselves, which is kept by their government. The court also refused to apply Article 10 to a case where safety precautions and information were not shared with the population surrounding a chemical plant. Protecting the health of others was similarly not enough of priority for preventing medical professionals from informing others of their methods in a public arena.

In protecting the general interest of the country, a teacher could not be fired due to his allegiance to a communist party, nor could obscene paintings be confiscated. However, a schoolbook dealing with sex in bad taste could be deemed to threaten the morals of the country. Individual rights are also a concern; a remark on the honesty of union representatives could not include the accusation of the individuals as thieves.

**ERROR ANALYSIS IDEAS**

Coefficients are given a weight with no consideration for fluctuation across time in the citation network. When case law becomes settled, do citation habits in the court change? LOOK UP IF ANY RESEARCH ON THIS. Do certain rights and freedoms enjoy a protection that has gone uncontested in recent decades, but whose early defense spawned considerable jurisprudence, and as a result stronger coefficients (and meme scores?) in a model trained on all years of a citation network statically.

Close look at meme article comparison methods. How many can I try

Memes on a more abstract level?

In methods express the thought given to considerations of your eval and their eval.

ERROR ANALYSIS IDEAS

Coefficients are given a weight with no consideration for fluctuation across time in the citation network. When case law becomes settled, do citation habits in the court change' LOOK UP IF ANY RESEARCH ON THIS. Do certain rights and freedoms enjoy a protection that has gone uncontested in recent decades, but whose early defense spawned considerable jurisprudence, and as a result stronger coefficients (and meme scores') in a model trained on all years of a citation network statically. \

Close look at meme article comparison methods. How many can I try

Memes on a more abstract level'

In methods express the thought given to considerations of your eval and their eval. like look over your code and why you did what

future research, context aware/paragraph to paragraph network, less noise, different sources cited too? like commission report? for rules/principles rate than just cases

s an illustration it might be worth pointing to numerous textbooks on the ECHR, which are structured around the convention text. While it is obvious that all cases are indeed adjudicated on the basis of the convention text (all cases are based on the violation of convention rights) it is not obvious that legal doctrine must be based only on these texts. As was pointed out by early legal realist theory, courts decide cases as much by way of responding to actual fact situations and background normativity as by way of applying a statutory (or treaty) text. See HERMAN OLIPHANT, A Return to Stare Decisis 14 American Bar Association Journal 71(1928). By arranging the case law according to the classification of the convention, doctrinal expositions may overlook fact patterns in cases that may be de facto “doctrinal” categories – for an example, MIRSHAHVALAD, et al., Significant Communities in Large Sparse Networks, 7 PLoS ONE e33721(2012)., which suggest that “borderline cases in the internal market” make up a separate category not previously recognized by scholarship (see further below)

many

traditional inferential techniques cannot be easily applied to network

data. The most popular statistical techniques rely on the assumption

that the data in question satisfies the IID assumption.93 However,

network data violates this assumption. Centrality measures for

instance are by definition not independent. One actor’s centrality in a

network depends on another actor’s centrality. Methodologists have

developed a number of network analytic techniques to help address

this issue.94 These often rely on simulating null models that

researchers can use to compare their observations against, in order to

determine the degree of “statistical significance” they can attribute to

their observations. One relatively straightforward approach is to use

quadratic assignment procedure (QAP) modeling, which permutes

the observed network a number of times in order to build a

probability distribution for the observed variables.95 This allows

researchers to compute standard errors and thus estimate statistical

significance.

Exponential Random Graph Models (ERGMs) are conceptually

similar, but much more powerful than a QAP analysis. ERGMs are a

family of statistical models developed for network analysis.96 They

rely on bayesian inference, and Markov Chain Monte Carlo

(MCMC) simulation to explain network structure as a function of

potentially related statistical parameters.97 This allows researchers to

determine the independent effect size (i.e., parameter estimates) for a

variety of network structures, and node or link-level attributes that

might contribute to the observed network’s structure.98 Applying

these sorts of advanced analytic techniques will help legal network

analyses move beyond description towards explanation, which will

represent a significant step in the progress of empirical legal studies