

Mapping Regulatory System of WTO Dispute Settlement Body Using Deep Learning

Suyeol Yun

November 13, 2020

Abstract

1 Introduction

The Dispute Settlement Body (DSB) of the World Trade Organization (WTO) deals with trade disputes between WTO members. WTO members can file a lawsuit in WTO DSB to claim their impaired benefit related to the WTO agreements as a result of possible illegal action of the other member's trade policy. Then a judicial body, *Panel* or *Appellate Body*, adjudicates the dispute and submits a report in which it expresses its judicial opinion as to whether the challenged trade policy is inconsistent to the rules of the WTO or not (World Trade Organization, 2017).

A lawsuit tends to cite multiple rules of the WTO agreement because one simple rule can't cover the complex characteristics of the trade policy that led to the dispute (Palmer and Mavroidis, 2004). For example, the United States enacted *Continued Dumping and Subsidy Act of 2000* (CDSOA) that distributes the collected anti-dumping duties to its affected domestic producers. This act was challenged by other members with multiple rules of the WTO agreements such as rules of *Anti-dumping* and rules of *Subsidy* because this distribution could constitute an illegal subsidy and anti-dumping duty at the same time (*See* Figure 1).

Citation of the rules of the WTO agreements tends to get more complicated because members cite the rules of the WTO agreements strategically. For example, members cite different rules of the WTO agreements to limit or to encourage the third party participation to the case. Since the third party participation can lead to early settlement of the dispute without continuous legal battle, members strategically cite different rules of WTO agreements according to their intention to settle the case earlier out of court or vice versa. (Johns and Pelc, 2014).

8.1 In the light of our findings, we conclude that **the CDSOA is inconsistent with AD (Anti-dumping) Articles 5.4, 18.1 and 18.4, SCM (Subsidy and Countervailing Measure) Articles 11.4, 32.1 and 32.5, Articles VI:2 and VI:3 of the GATT 1994, and Article XVI:4 of the WTO Agreement.**

...

8.3 **The CDSOA is a new and complex measure, applied in a complex legal environment.** In concluding that the CDSOA is in violation of the above mentioned provisions, we have been confronted by sensitive issues regarding the use of subsidies as trade remedies. this matter through negotiation.

Figure 1: **Panel’s Judicial Opinion On the *US - Offset (Byrd Amendment; CDSOA)* case:** Panel explicitly expresses the complexity of the trade policy (CDSOA) at issue and cites the rules of anti-dumpig (AD) and subsidy (SCM) at the same time to cover its complex characteristics.

In addition to it, members’ citation of articles of the WTO agreements also gets complicated if we consider the fact that *Panel* or *Appellate Body* defers to the legal precedents of WTO DSB. Legal precedents refer to *Panel* or *Appellate Body*’s judicial decisions and these legal precedents provide authoritative reference for deciding subsequent cases in WTO DSB. Members try to re-shape these legal precedents in favor of their future interest more than simply using the WTO DSB to resolve their trade dispute (Pelc, 2014). For example, members tend to cite their favorable previous cases more often in issue areas where they face litigation more frequently with other members (Strezhnev, 2014).

As WTO sets its main principles to regulate the world trade system in general such as principle of *Market Access* (across borders), *Non-discrimination* (between members or between domestic products and imported products) and *Transparency* (in publication and maintaining of each member’s internal regulations), it’s intellectually intriguing to understand how regulatory system of WTO DSB is structured to achieve these main principles (*See* Figure 2). By understanding this structure, we can improve WTO system to serve its main principles more effectively and to adopt to constantly changing world trade circumstances (Fredebeul-Krein and Freytag, 1999; Shaffer, 2004; Hufbauer and Stephenson, 2007).

However, it is extremely difficult to understand how rules of the WTO agreements are working together to achieve those main principles of WTO. This is because each citation is closely related to complex characteristics of each trade policy as exemplified in the *US - Offset (Byrd Amendment; CDOSA)* case (Figure 1). Moreover, to understand interactions between multiple rules of WTO agreements properly, it requires to generalize members’ strategic citation patterns which are

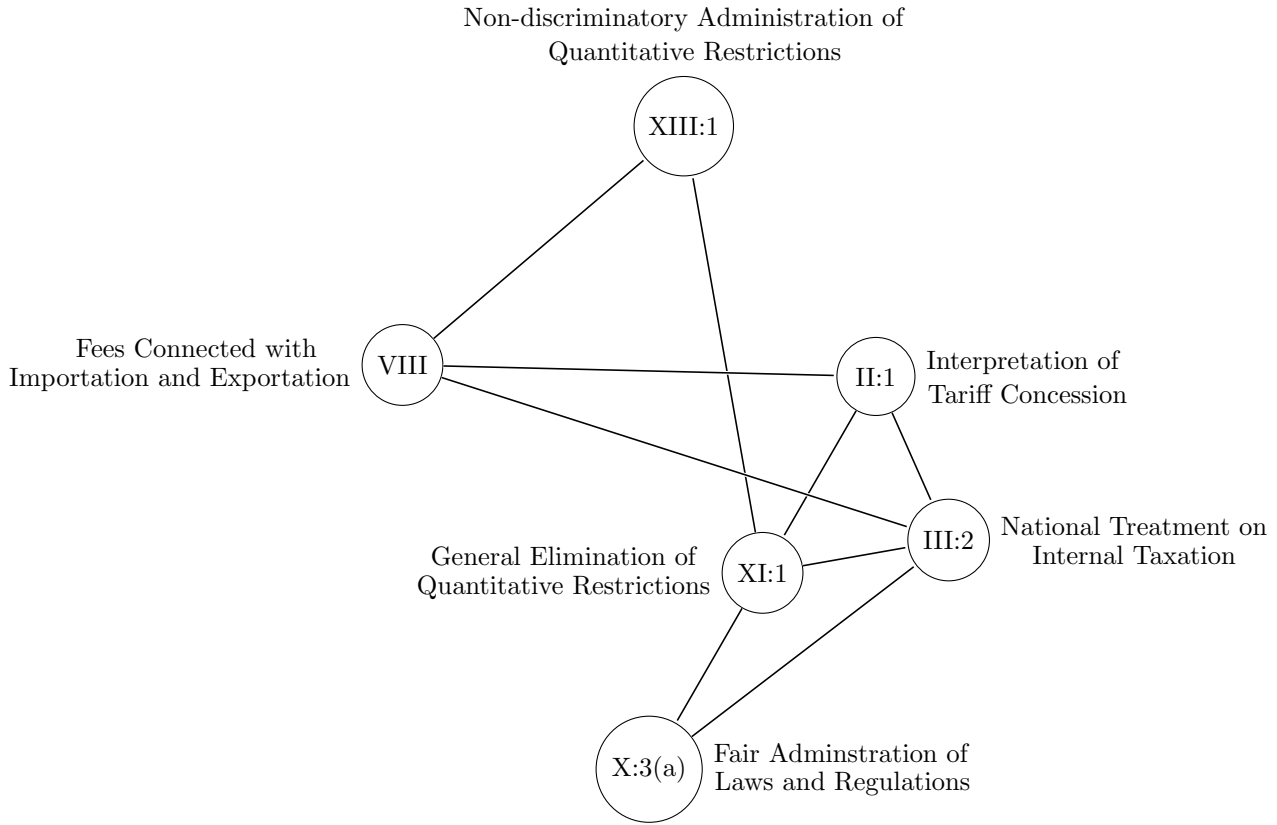


Figure 2: **Network of the Articles that Achieves *Market Access* Together:** This figure demonstrates a network of articles of WTO agreements that cooperatively achieves the principle of *Market Access* in the regulatory system of WTO. Tariff and Non-tariff barriers such as quantitative restriction, internal taxations and extra fees for crossing border can inhibit the chance of foreign goods to access the foreign market. Therefore, these articles tend to work together to ensure the *Market Access* principle working properly.

limited to each member's special interest rather than explaining the regulatory system of WTO DSB in general.

To address this issue, this paper maps the regulatory system of WTO DSB as a network of legal articles of the WTO agreements as formally defined in Figure 3 and illustrated in Figure 4. This is because the rules of the WTO agreements explicitly requires *Panel* or *Appellate Body* to address relevant articles together when they construct its jurisprudence related to the meaning, scope and interpretation of any legal text in the WTO agreements (*See* Figure 5). Upon this requirement, judicial bodies refer to multiple articles of the WTO agreements together to identify the complex legal identity of the trade policy that led to the dispute as shown in Figure 1. In addition to it, judicial bodies cite multiple articles together to guide an way of interpretation of the rules of the WTO agreements (*See* Figure 4(b))

To map the regulatory system of WTO DSB as a network of legal articles of WTO agreement

Network of legal articles of WTO agreements is defined as

$$\text{directed weighted graph } G = (V, E, W)$$

where *vertex set* $V = \{v \mid v \text{ is a legal article of WTO agreement}\}$,

set of directed edges $\vec{E} = \{(v_i, v_j) \mid (v_i, v_j) \in V \times V\}$ and

$$w : V \times V \rightarrow \mathbb{R}_+ \text{ s.t. } w(v_j, v_j) = 0 \text{ and } \sum_{v_i \in V} w(v_i, v_j) = 1 \forall v_j \in V$$

Then define *edge weight matrix* $W = (w_{ij}) \in \mathbb{R}_+^{|V| \times |V|}$ s.t. $w_{ij} = w(v_i, v_j)$

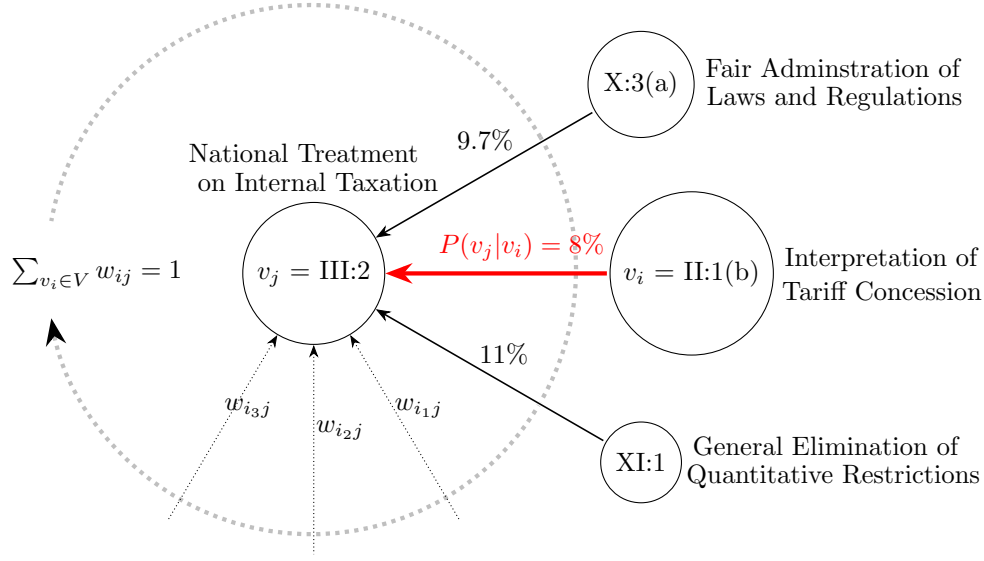
(W is more formally called *weighted adjacency matrix*)

Figure 3: **Formal Definition of Network of Legal Articles of WTO agreements:** I define network of legal articles of WTO agreements as a directed weighted graph where the sum of all weights coming into a node sum up to 1. w_{ij} is interpreted as conditional probability $P(v_j|v_i)$ how probably a source node v_i clarifies the meaning of the target node v_j compared to other source nodes as illustrated in Figure 4

successfully, this paper designs a deep neural network (Figure ??) that processes two different types of textual information. One is textual description of the dispute (See an example at Appendix A.1) and the other one is the text of a legal article of the WTO agreements (See an example at Figure 6). This design is improvised to mimic the reasoning process of WTO legal practitioners where the legal practitioners read the textual description of factual circumstances of the dispute and imagine applicable regulatory contents of the legal articles while he/she reads the factual description (See 15, 16 and 17).

To train this neural network, this paper collected textual description of trade policy that led to the dispute and articles of the WTO agreement cited for each dispute case requested to the WTO DSB from 1995 to 2018 (Total 143 cases. Check the list in Appendix A.2). Using this collected data, I trained the neural network by enforcing the neural network to answer correctly whether a given article of the WTO agreements can be cited for the given textual description of trade policy that led to the dispute (See Figure 17 and Figure 18). After training, I fitted a set of directed edge weight W^* that best explains the variance of each article's citability predicted by the trained deep neural network using an ensemble of random forests (Huynh-Thu et al., 2010).

To check whether this fitted network of articles of the WTO agreements $G^* = (V, E, W^*)$ maps the regulatory system of WTO DSB properly, this paper compares the created network and the jurisprudence of WTO DSB made by *Panel* and *Appellate Body*. This comparison reveals that the fitted network G^* captures the interaction between the articles of WTO agreements similarly with the jurisprudence of *Panel* and *the Appellate Body*. This similarity guarantees that the fitted



(a) Illustrated edge weights of a source node Article II:1(b)

“The dictionary definition of the noun ‘excess’ is ‘[t]he amount by which one number or quantity exceeds another’. More specifically, ‘in excess of’ means ‘more than’. Thus, as a textual matter, a particular number or quantity is ‘in excess of’ another number or quantity if it is greater, regardless of the extent to which it is greater. ***Looking at the context of Article II:1(b), first sentence, we note that Article III:2, first sentence, of the GATT 1994 is cast in very similar terms and in fact uses the phrase ‘in excess of’:***

The products of the territory of any contracting party imported into the territory of any other contracting party shall not be subject ... to internal taxes or other internal charges of any kind in excess of those applied ... to like domestic products ...

(b) **Jurisprudence of Panel in *Russia – Tariff Treatment* case:**

Panel clarifies the point that the meaning of the term ‘*in excess of*’ in Article II:1(b) clarifies the meaning of the same phrase in Article III:2.

Figure 4: **Illustration of Network of Legal Articles of WTO agreements:** Every directed edge weight w_{ij} is interpreted as the conditional probability $P(v_j|v_i)$ of how probably a source node v_i constitutes a legal context to clarify the meaning of the target node v_j among all target nodes $v_{j \neq i} \in V$. Above subfigure (a) represents how jurisprudence of *Panel* stated in (b) is represented as an edge weight 8% where the source node Article II:1(b) constitutes the legal context of the target node Article III:2 with the probability of 8% compared to other source articles.

Article 7

Terms of Reference of Panels

1. Panels shall have the following terms of reference unless the parties to the dispute agree otherwise within 20 days from the establishment of the panel:

“To examine, **in the light of the relevant provisions** in (name of the covered agreement(s) cited by the parties to the dispute), the matter referred to the DSB by (name of party) in document . . . and to make such findings as will assist the DSB in making the recommendations or in giving the rulings provided for in that/those agreement(s).”

2. **Panels shall address the relevant provisions** in any covered agreement or agreements cited by the parties to the dispute. . . .

Figure 5: **Article 7 of the Dispute Settlement Understanding (DSU):** DSU provides a legal guidelines on how judicial bodies of WTO shall adjudicate the requested disputes. It explicitly requires judicial bodies to interweave relevant articles of the WTO agreements to clarify its meaning, scope and interpretation.

network G^* closely maps the regulatory system of WTO DSB since only these two judicial bodies can authoritatively constitute the jurisprudence over how rules of WTO agreements are working together to achieve the main principles of WTO.

Moreover, this paper justifies the use of textual information inside the textual description of the dispute and legal article by showing that simply using the co-citation pattern between articles of the WTO disputes can't qualitatively fit the G^* (See Figure ?? and Figure ??). Upon this necessity of using the textual information, this paper also justifies the use of neural network that is computationally intensive since it's generally known that proper design of neural network is able to effectively extract information from the textual content.

Finally, I will explain how this fitting of network of legal articles can contribute to the current study of international normative system.

2 Data: Types, Composition and Collection Process

This section explains the composition of data and its collection process in detail.

General Most-Favoured-Nation Treatment

1. With respect to customs duties and charges of any kind imposed on or in connection with importation or exportation or imposed on the international transfer of payments for imports or exports, and with respect to the method of levying such duties and charges, and with respect to all rules and formalities in connection with importation and exportation, and with respect to all matters referred to in paragraphs 2 and 4 of Article III, any advantage, favour, privilege or immunity granted by any contracting party to any product originating in or destined for any other country shall be accorded immediately and unconditionally to the like product originating in or destined for the territories of all other contracting parties...

Figure 6: **Example of content of a legal article of the WTO agreement:** Article I:1 of GATT 1994 that prohibits the discrimination among similar products (WTO DSB prefer to call *like products*).

2.1 Overview: How Members Raise Claims in WTO DSB

As explained in the introduction, a trade policy that led to a dispute (preferably called as *Government Measure* in WTO DSB) is pretty much complicated as explicitly expressed by the Panel in Figure 1.

To address this complexity, members who raise the claim (preferably called *complainant* in WTO DSB) usually cite multiple articles of the WTO agreements at the same time. For example, in the *US - Offset (Byrd Amendment)* case, a group of complainants cited articles as shown in Table 1 from the WTO agreements to claim the inconsistency of *Continued Dumping and Subsidy Act of 2000* (CDSOA) of United States to these articles¹:

Name of WTO Agreement	Cited Articles
Agreement on Anti-dumping	1, 5.4, 8, 18.1, 18.4
General Agreement on Tariffs and Trade 1994	VI:3, X:3, XXIII:1, VI:2
Agreement on Subsidies and Countervailing Measures	4.10, 7.9, 10, 11.4, 18, 32.1, 32.5
Agreement Establishing the World Trade Organization	XVI:4

Table 1: Cited articles in *US - Offset (Byrd Amendment)* by complainants

Upon this understanding, I collected two different types of data for 143 different dispute cases requested to WTO DSB. (List of cases is available at Appendix A.2). One is textual descrip-

⁰Australia, Brazil, Chile, European Communities, India, Indonesia, Japan, Korea and Thailand

¹It is worth noting that the WTO agreements comprises many different agreements covering each specific topic in trade such as *Agreement on Anti-dumping*, *Agreement on Subsidies and Countervailing Measures*, *Agreement on Agriculture* and so on.

tion of the dispute (*Check* the CDSOA example at Appendix A.1) and the other one is set of articles of the WTO agreements that are cited for each dispute (Appendix A.3). I will explain source, structure and collection method for two different types of data at the following subsections.

		WT/DS217/R WT/DS234/R Page i
TABLE OF CONTENTS		
		<u>Page</u>
I.	INTRODUCTION	1
II.	FACTUAL ASPECTS	2
III.	PARTIES' REQUESTS FOR FINDINGS AND RECOMMENDATIONS	4
A.	COMPLAINING PARTIES	4
B.	UNITED STATES	5
IV.	ARGUMENTS OF THE PARTIES	5
A.	FIRST WRITTEN SUBMISSION OF THE COMPLAINING PARTIES.....	6
1.	Australia	6
2.	Brazil	9
3.	Canada	15
4.	Chile and Japan	24
5.	European Communities, India, Indonesia and Thailand	32
6.	Korea	40
7.	Mexico	45

Figure 7: **Table of Contents of Panel Report:** Panel provides factual aspect in the panel report with its page location.

2.2 Factual Aspect: Textual Description of the Dispute

Textual description of the dispute is preferably called as *Factual Aspect* in WTO DSB. Since Panels always provide a factual aspect² that summarizes the content of the dispute in the panel report, I wrote a program that can automatically search and collect the panel reports from the WTO official document website³. Then I located the factual aspect using the page information inside the table of contents in the panel report as shown in Figure 7. By using this location, I excerpted factual aspect from 143 number of different panel reports listed in Figure 8.

²It's worth noting that Appellate Body doesn't provide any factual aspect because they use the factual aspect provided by the Panel.

³<http://docs.wto.org>

2.2.1 Joint Adjudication & Early Settlement

The number 143 seems small compared to the total number (596⁴) of cases requested to WTO DSB.

This is because, first, panel handles different cases together if the case is about the same trade policy. For example, in *US - Offset (Byrd Amendment)*, panel merged DS217⁵ and DS234 together because they were asking the judicial opinion for the same government measure of the United States as shown in Figure 9. This paper selects the smallest case number as a representative number for this kind of joint Adjudication. For example, since DS217 and DS234 shares the same panel report, this paper choose DS217 as a representative number as shown in Figure 8 where the list includes DS217 but not DS234.

Secondly, members sometimes find *mutually agreeable solution* before the panel expresses its judicial opinion by publishing its panel report. Then Panel stops there and no factual aspect is available. This paper omitted this kind of early settled cases as well.

DS 2, 18, 22, 31, 34, 46, 56, 58, 60, 62, 67, 68, 69, 75, 76, 87, 90, 98, 103, 108, 121, 122, 135, 136, 139, 141, 146, 152, 155, 161, 162, 165, 166, 174, 175, 177, 184, 202, 207, 212, 217, 219, 221, 231, 234, 238, 244, 245, 246, 248, 257, 264, 265, 266, 267, 268, 269, 276, 282, 283, 286, 290, 294, 295, 296, 301, 302, 308, 312, 315, 316, 320, 321, 322, 332, 336, 339, 343, 344, 345, 350, 353, 360, 363, 366, 371, 379, 381, 384, 392, 394, 396, 397, 399, 400, 406, 412, 414, 415, 422, 425, 427, 429, 430, 431, 435, 436, 437, 440, 442, 447, 449, 453, 454, 456, 457, 461, 464, 468, 471, 472, 473, 475, 476, 477, 479, 480, 482, 483, 484, 485, 486, 488, 490, 492, 493, 495, 499, 504, 505, 513, 518, 523

Figure 8: **List of case number of collected panel reports:** “DS + number” uniquely identifies each dispute. For example, DS 523 refers to *US — Pipe and Tube Products (Turkey)* where United States was challenged by Turkey for its possibly inconsistent anti-dumping measure.

2.3 Cited Articles: Set of Articles Cited for the Same Dispute

Every lawsuit in WTO DSB cites multiple set of articles as shown in Table 1. To collect this set of articles claimed for the same dispute, I wrote a program that collects this set of articles cited for the same dispute from the WTO official webpage⁶. The webpage chronologically lists up all dispute cases requested to WTO DSB and the program visits each page of 143 cases and collects

⁴As of November 1st, 2020.

⁵DS refers to Dispute Settlement. DS is official prefix that indicates the case in WTO DSB.

⁶https://www.wto.org/english/tratop_e/dispu_e/dispu_status_e.htm

**UNITED STATES – CONTINUED DUMPING AND SUBSIDY
OFFSET ACT OF 2000**

Report of the Panel

Figure 9: **Cover of Panel Report:** Panel explicitly marks which different cases are handled together in the cover of the panel report. DS217 and DS234 are handled together in this example.

the cited articles. Among all the articles from different agreements of the WTO agreements⁷, this paper collected articles from **General Agreement on Tariffs and Trade 1994 (GATT 1994)** only. This is because articles in GATT 1994 constitutes basic set of trade rules of WTO and other agreements elaborates the articles of GATT 1994 more in detail (World Trade Organization, 1999). For example, the official name of *Agreement on Anti-dumping* is *Agreement on Implementation of Article VI of the GATT 1994* where the name self-explains that it elaborates on the article VI of GATT 1994. The collected result is listed in the Appendix A.2. Figure 10 lists up 80 different articles of GATT 1994 cited in 143 cases without duplication.

2.3.1 Various Levels of Scope in Cited Articles

As shown in Figure 10, members sometimes cite articles in different levels of scope. For example, For the Article VI, member sometimes cites Article VI as a whole but sometimes cites Article VI:2 or Article VI:2(a). This is because two main judicial bodies of WTO DSB, *Panel and Appellate Body*, both constitute its legal precedents citing articles of the WTO agreements in various levels

⁷WTO agreements is comprised of multiple agreements such as General Agreement on Tariffs and Trade 1994, Agreement on Agriculture, Agreement on the Application of Sanitary and Phytosanitary Measures, Agreement on Textiles and Clothing, Agreement on Technical Barriers to Trade, Agreement on Trade-Related Investment Measures, Agreement on Implementation of Article VI of the General Agreement on Tariffs and Trade 1994 (anti-dumping), Agreement on Subsidies and Countervailing Measures, Agreement on Rules of Origin, Agreement on Safeguards and so on.

I, I:1, II, II:1, II:1(a), II:1(b), II:2, II:3, III, III:1, III:2, III:4, III:5, III:7, IV, IX, IX:2, V, V:1, V:2, V:3, V:3(a), V:4, V:5, V:6, V:7, VI, VI:1, VI:2, VI:2(a), VI:2(b), VI:3, VI:5(a), VI:6, VII, VII:1, VII:2, VII:5, VIII, VIII:1, VIII:3, VIII:4, X, X:1, X:2, X:3, X:3(a), XI, XI:1, XIII, XIII:1, XIII:2, XIII:3(b), XIX, XIX:1, XIX:2, XIX:3, XV, XVI, XVI:1, XVI:4, XVII, XVII:1, XVII:1(c), XVIII, XVIII:10, XVIII:11, XX, XXI, XXII, XXII:1, XXIII, XXIII:1, XXIII:1(a), XXIII:1(b), XXIV, XXIV:12, XXIV:5(b), XXIV:6, XXVIII

Figure 10: **Set of articles of GATT 1994 collected and used in this paper:** These articles comprises the node set V and their ordered pairs comprise the edge set E in Figure 3

Let D is a set of DS case numbers listed in Figure 8.

Then there exists $c_d = \{v_d \in V \mid v_d \text{ is an article cited in the case } d\} \forall d \in D$

where V is set of articles listed in Figure 10.

Then define set of cited articles $C = \{c_d \mid d \in D\}$

Figure 11: **Formal Definition of Set of Cited Articles:** I formally define a set of cited articles C and the elements of C are listed in Appendix A.2.

of scope. Both judicial bodies cite the legal articles with the level of *Title*, *Article*, *Paragraph*, *Sentence* or *Term* as shown in Table 2. Following this jurisprudence, members also cite articles in different levels of scope to make their legal claim fit and valid according to the current jurisprudence of WTO DSB.

Table 2: **Various Levels of Scope Adopted to Cite Articles of WTO agreemnts**

Scope	Quote	Source
Title	“As the <i>title</i> of Article 21 makes clear , the task of panels . . . forms part of the process of the ‘Surveillance of Implementation of the Recommendations and Rulings’ of the DSB. . . .”	Appellate Body Report, <i>US – Shrimp (Malaysia)</i> , paras. 86-87.

Article	“The sequence of steps indicated above in the analysis of a claim of justification under Article XX reflects, not inadvertence or random choice, but rather the fundamental structure and logic of Article XX. ...”	Appellate Body Report, <i>US – Shrimp (Malaysia)</i> , paras. 119-120.
Paragraph	“The verb ‘may’ in Article VI:2 of the GATT 1994 is, in our opinion, properly understood as giving Members a choice between imposing an anti-dumping duty or not, as well as a choice between imposing an anti-dumping duty equal to the dumping margin or imposing a lower duty. ...”	Appellate Body Report, <i>US – 1916 Act</i> , paras. 116.
Sentence	“The customary rules of interpretation of public international law as required by the first sentence of Article 17.6(ii) of the Anti-Dumping Agreement , do not admit of another interpretation as far as the issue of zeroing raised in this appeal is concerned.”	Appellate Body Report, <i>US – Zeroing (EC)</i> , paras. 132-133.
Term	“Article II:1(a) provides that a Member shall accord to the ‘commerce’ of other Members treatment no less favourable than that provided for in its Schedule. The term ‘commerce’ is defined as referring broadly to the exchange of goods such that, in this provision, the ‘commerce’ of a Member should be understood to refer to all such exchanges of that Member”	Appellate Body Report, <i>Colombia – Textiles</i> , para. 5.34.

3 Methodology: Considerations and Development

This section introduces two main considerations to design the method used in this paper. Then it explains the method that is used to fit the network of articles of WTO agreement under those considerations.

3.1 Two Main Considerations For Design of Method

This paper considered two main points to determine its method to qualitatively fit a set of edge weight W for the *directed weighted graph* G defined in Figure 3. One is importance of using the information represented in a form of textual description inside the content of dispute and legal article as exemplified in Appendix A.1 and Figure 6 respectively. The other one is about the way to generalize each member’s strategic citation pattern. Since members of the WTO strategically cite the articles of WTO agreement expecting different outcomes that serves member-specific national interest (Johns and Pelc, 2014; Pelc, 2014; Strezhnev, 2014), this paper selected a method that can generalize this member specific citation pattern. These two considerations and the solution will be explained in the following subsections.

3.1.1 Importance of Using Textual Information

This paper emphasizes the necessity of using textual information to qualitatively fit a set of edge weight W defined in Figure 3. One can simply consider a co-citation pattern between the articles of WTO agreements as a regulatory system of WTO DSB, however, it simply allocates a large edge weight for frequently cited articles and fails to explain how articles interact to achieve main principles of WTO not like well exemplified in Figure 2.

This failure is mainly due to the insufficient information in co-citation matrix. Members tend to cite the articles of the WTO agreements based on the complex characteristics of the trade policy that led to the dispute, however, co-citation pattern omits this prior information.

To emphasize the importance of using textual information, I prepared two different matrices $W_{\text{co-cites}}$ and W_{text} that is following the definition of *edge weight matrix* W in Figure 3. $W_{\text{co-cites}}$ is calculated only using the co-citation pattern between the articles of the WTO agreements as formally defined as *Normalized Co-citation Matrix* in Figure 14. W_{text} is the one that is fitted using the textual information and the way how it’s fitted will be introduced at the following body in this section. Two Heatmaps visualized in Figure 12 shows how sparse the $W_{\text{co-cites}}$ is. This sparsity indirectly refers to the insufficient information to qualitatively map the regulatory system

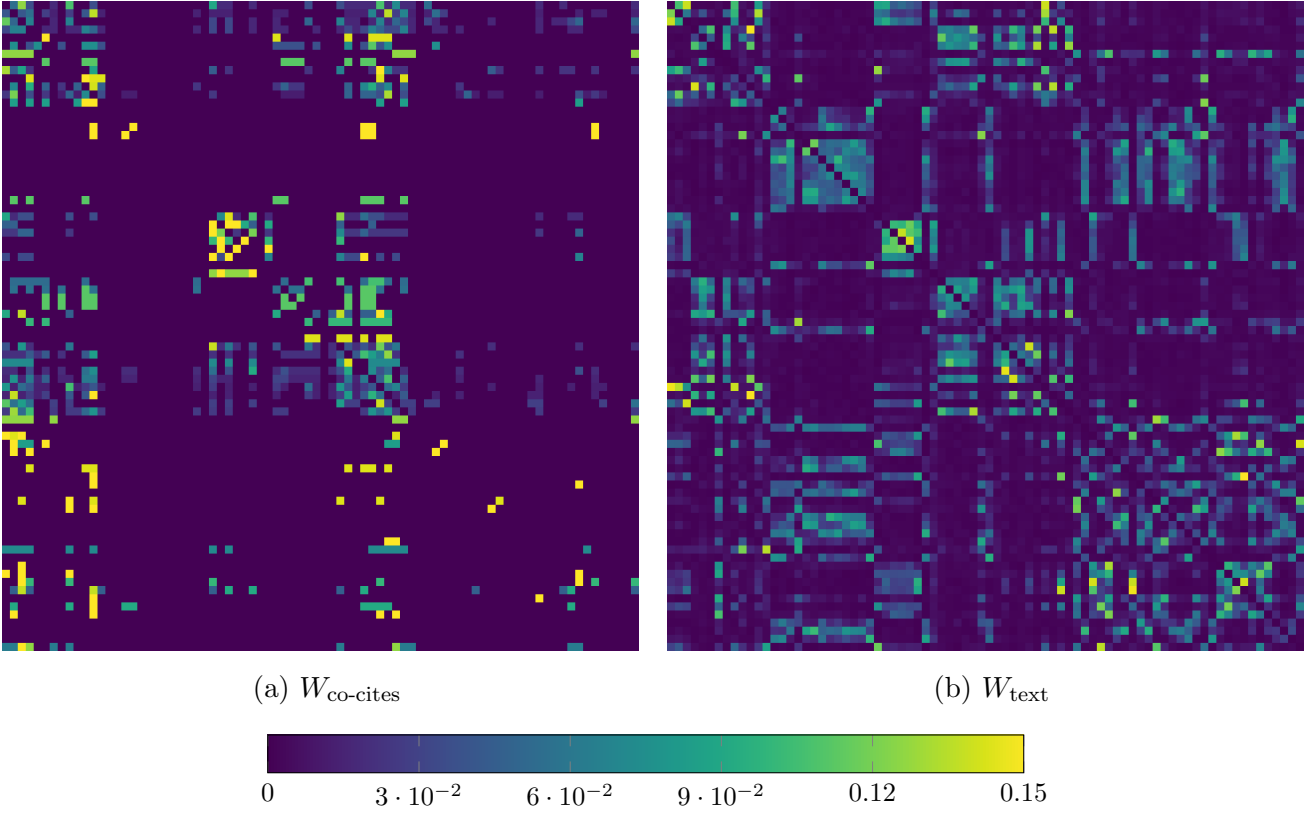


Figure 12: **Heatmap of Two Different Edge Weight Matrices:** Above two subfigures visualizes two different *edge weight matrices* $W_{\text{co-cites}}$ and W_{text} . One can check that $W_{\text{co-cites}}$ is sparser than W_{text} .

of WTO DSB. In contrast with it, if we fit the *edge weight matrix* W using the textual information, we get more dense matrix as visualized in Figure12(b).

Upon this observation, this paper justifies the use a deep neural network to process information embedded in the text description of the trade policy and legal articles. This is because deep neural network is known to extract information from the text effectively to perform various tasks such as text classification (Minaee et al., 2020), text summarization (Magdum and Rathi, 2021) and text generation (Guo et al., 2017).

3.1.2 Generalization of Each Member’s Strategic Citation

This paper aims to map the regulatory system of WTO DSB in a form of *directed weighted graph* G as defined in Figure 3. To achieve this purpose, we need to fit W to generalize member specific strategic citation behavior. In terms of generalization, deep neural network is known to generalize well despite its large capacity (Neyshabur et al., 2017), possible instability of training algorithm (Charles and Papailiopoulos, 2017), nonrobustness (Zahavy et al., 2017), and sharp minima (Dinh et al., 2017). Therefore this paper trains a deep neural network without any member specific

Let δ_{ij}^d is defined to be 1 if $\{(v_i, v_j) \mid v_i, v_j \in V \text{ and } i \neq j\} \subset c_{d \in D}$ else 0

where V, D and c_d is defined as in Figure 11.

Then let *co-citation matrix* $M = (m_{ij}) \in \mathbb{N}^{|V| \times |V|}$ s.t. $m_{ij} = \sum_{d \in D} \sum_{i, j \in V} \delta_{ij}^d$

(a) **Formal Definition of Co-citation Matrix**

	I	I:1	II	II:1	...
I	0	3	7	2	
I:1	3	0	3	4	
II	7	3	0	4	
II:1	2	4	4	0	
⋮					

(b) **Illustration of Co-citation Matrix**

Figure 13: **Formal Definition and Illustration of Co-citation Matrix:** This paper defines co-citation matrix M as subfigure (a) and it's illustrated as subfigure (b) using the paper's dataset. Note that co-citation matrix is *symmetric*, $m_{ij} = m_{ji} \forall i, j \in V$.

information such as geolocation, GDP or specialized industry and so on. By training a deep neural network only using a description of the possible inconsistent trade policy and the legal text of the cited articles, this paper expects a fitted $G^* = (V, E, W^*)$ can show interactions between articles of the WTO agreements without being biased to member specific strategic citation.

3.2 Design of Deep Neural Network

Upon the justification of using deep neural network with two main reasons explained in Section 3.1, I introduce a design of deep neural network that is expected to efficiently encode members' legal citation pattern.

3.2.1 Design Input/Output of Deep Neural Network: by Analogy with How Member Cites in WTO DSB

A rule of thumb to design input and output of deep neural network is to mimic how humans do for a given task. Therefore I present a visualization of how legal expert of rules of WTO determines whether to cite a legal article of WTO agreements with an example of *Korea - Beef*

For given M defined in Figure 13(a),

let *normalized co-citation matrix* $N = (n_{ij}) \in \mathbb{R}^{|V| \times |V|}$ s.t. $n_{ij} = \frac{m_{ij}}{\sum_{j \in V} m_{ij}}$

(a) **Formal Definition of Normalized Co-citation Matrix**

	I	I:1	II	II:1	...
I	0	0.053	0.125	0.035 $\rightarrow \sum_{j \in V} n_{ij} = 1$
I:1	0.040	0	0.04	0.054	
II	0.114	0.049	0	0.065	
II:1	0.032	0.065	0.065	0	
\vdots					

(b) **Illustration of Noramlized Co-citation Matrix**

Figure 14: **Formal Definition and Illustration of Normalized Co-citation Matrix:** This paper defines normalized co-citation matrix N of M as subfigure (a) and it's illustrated as subfigure (b) using the paper's dataset. Note that normalized co-citation matrix is no more *symmetric*, $n_{ij} \neq n_{ji} \forall i, j \in V$. This definition is prepared to fit the definition of co-citation matrix to that of W in Figure 3.

case (Figure 15 and Figure 16). In this case, the United States raised a claim relating to the *Dual-retail system* maintained by South Korea. In *Dual-retail system*, South Korea maintained two distinct retail systems for imported and domestic beef, i.e., there exist stores specialized for imported beef and they can sell only imported beef and cannot sell domestic beef. U.S. claimed that the *Dual-retail system* is inconsistent to the Article III:4 (National Treatment) of GATT 1994 because *Dual-retail system* has a chance to discriminate domestic and imported beef. A measure that discriminates domestic and imported products falls under the scope of the Article III:4 of the GATT 1994, which states the principle of *National Treatment* that prohibits the discrimination between imported and domestic *like product*. However, U.S. didn't cite the Article I:1 of GATT 1994 that prohibits the discrimination between members of WTO agreements because *Dual-retail system* didn't discriminate the United States with other countries that exports beef to South Korea, such as Argentina, Austrailia, etc.

We can understand that there exists a shared understanding over regulatory system of WTO DSB among legal experts of the rules of WTO. They follow up new cases and study jurisprudence stated in the *Panel* or *Appellate Body* reports. Then they organize a legal argument by citing

certain article(s) upon this shared understanding for given possible inconsistent measure held by a member of WTO.

To mimic this reasoning process, I designed the input and output of the deep neural network as illustrated in Figure 17. The neural network is designed to estimate the citable probability for given pair of a factual aspect and text of a legal article. By iteratively trains the neural network with data explained in Section 2, I expect the neural network can learn a shared understanding over regulatory system of WTO DSB as closely as to that of legal experts. The detailed structure of neural network and training technique will be explained in the later subsections.

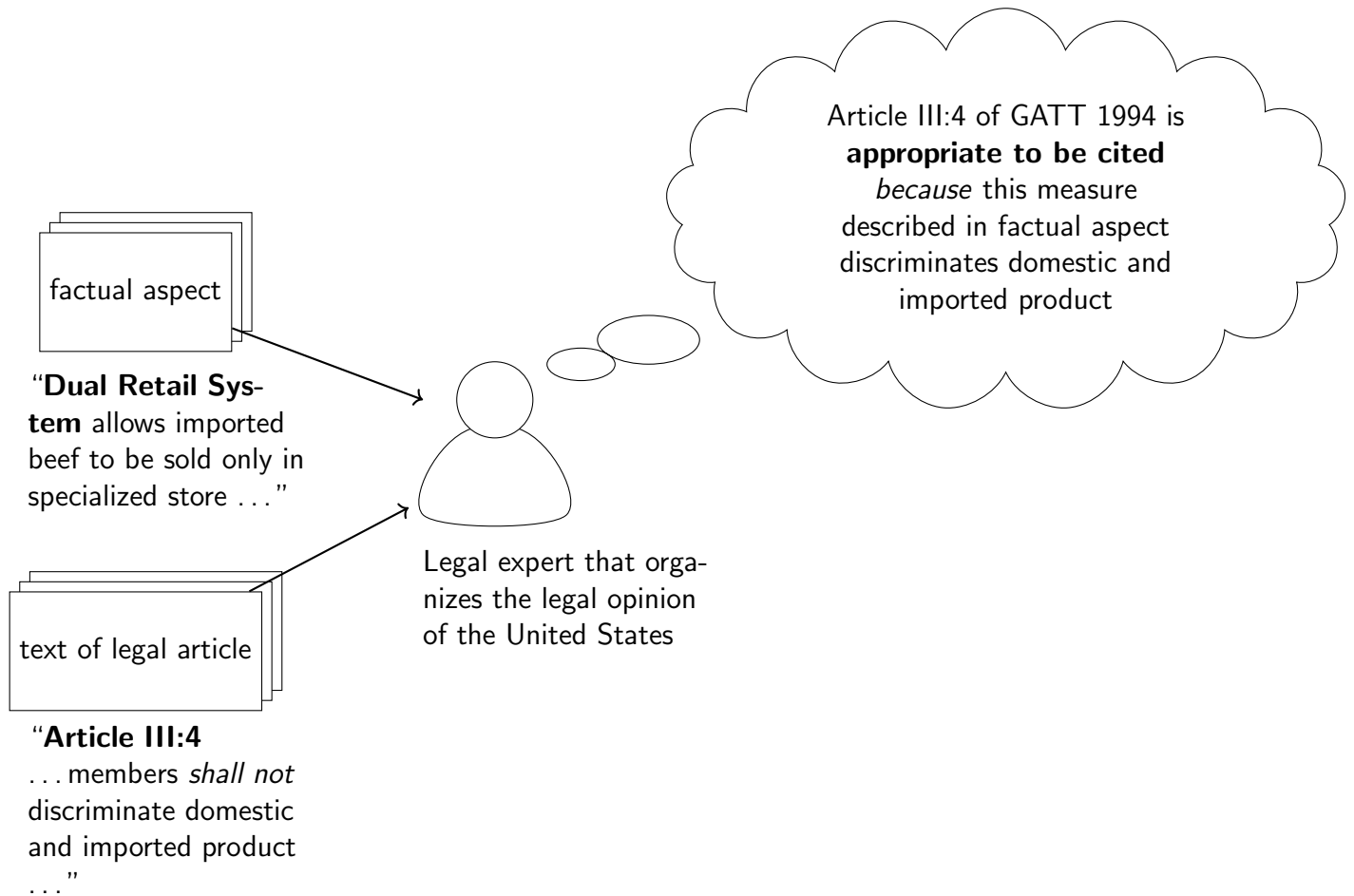


Figure 15: **Visualization of How Member Cites in WTO DSB (Citable Case):** With two different contexts, factual aspect and text of legal articles, member judges whether the given legal article is appropriate to be cited or not.

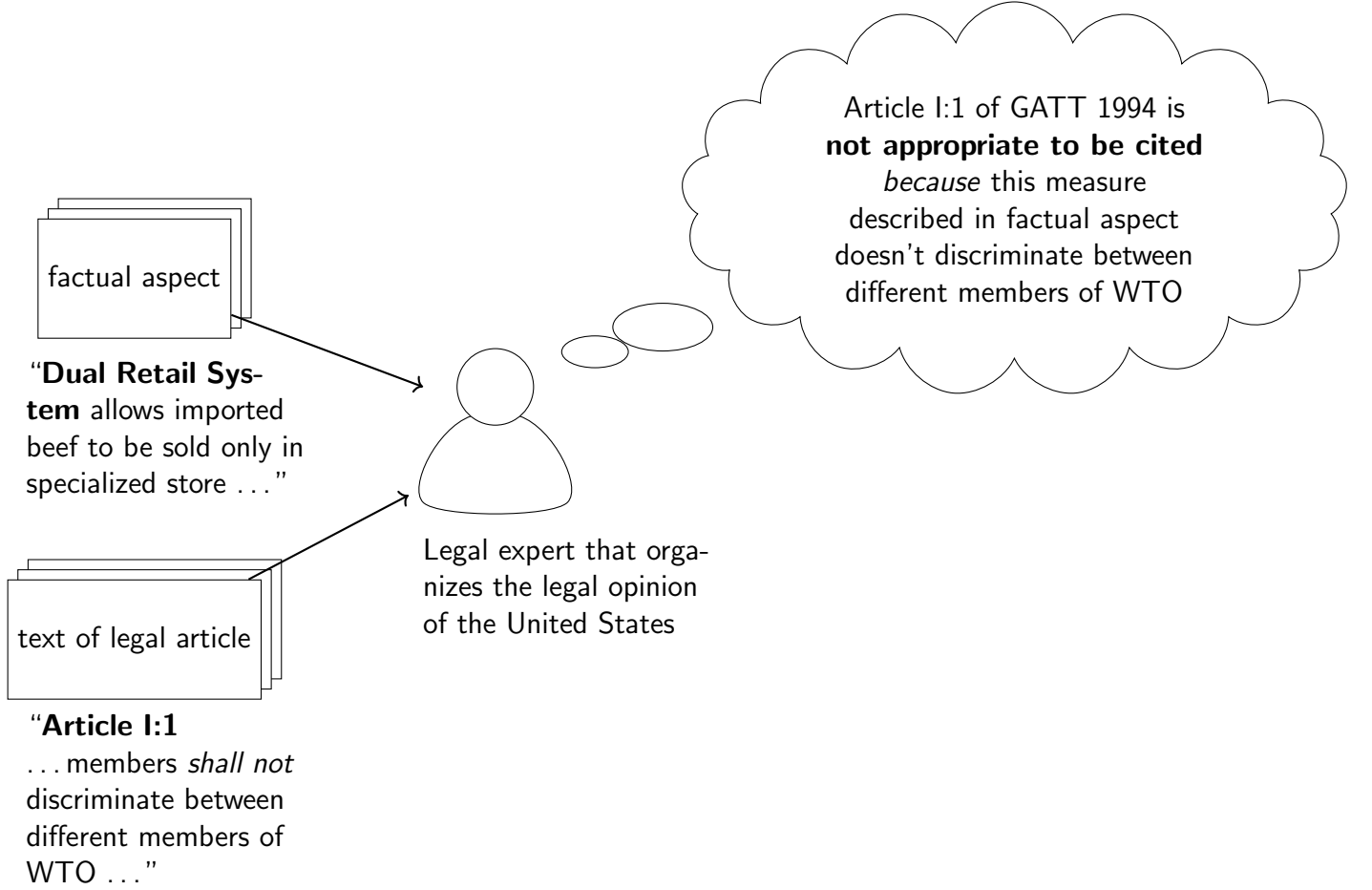


Figure 16: **Visualization of How Member Cites in WTO DSB (Non-Citable Case):** With two different contexts, factual aspect and text of legal articles, member judges whether the given legal article is appropriate to be cited or not.

3.2.2 Structure of Deep Neural Network

3.2.3 Train of Deep Neural Network

3.3 Fitting $G^* = (V, E, W^*)$ Using Ensemble of Random Forests

3.3.1 Definition of W^* : Best Set of Directed Edge Weights

Let f_{θ^*} represent the trained neural network with an optimized set of parameters θ^* . Then we can construct a prediction matrix $P = (f_{\theta^*}(t_d, a_v)) \in [0, 1]^{|D| \times |V|}$ by collecting all predictions $f_{\theta^*}(t_d, a_v)$ from the trained neural network f_{θ^*} using the all pairs of data $(t_d, a_v) \in T \times A$ as illustrated in Figure 22.

Upon the assumption that the trained neural network f_{θ^*} effectively encodes a shared understanding among legal experts of WTO agreements, our task is to find a set of directed edge weights $\{w_{ij} \mid w_{ij} = w(v_i, v_j) \text{ s.t. } w(v_j, v_j) = 0 \text{ and } \sum_{v_i \in V} w(v_i, v_j) = 1 \forall v_j\}$ as defined in Figure

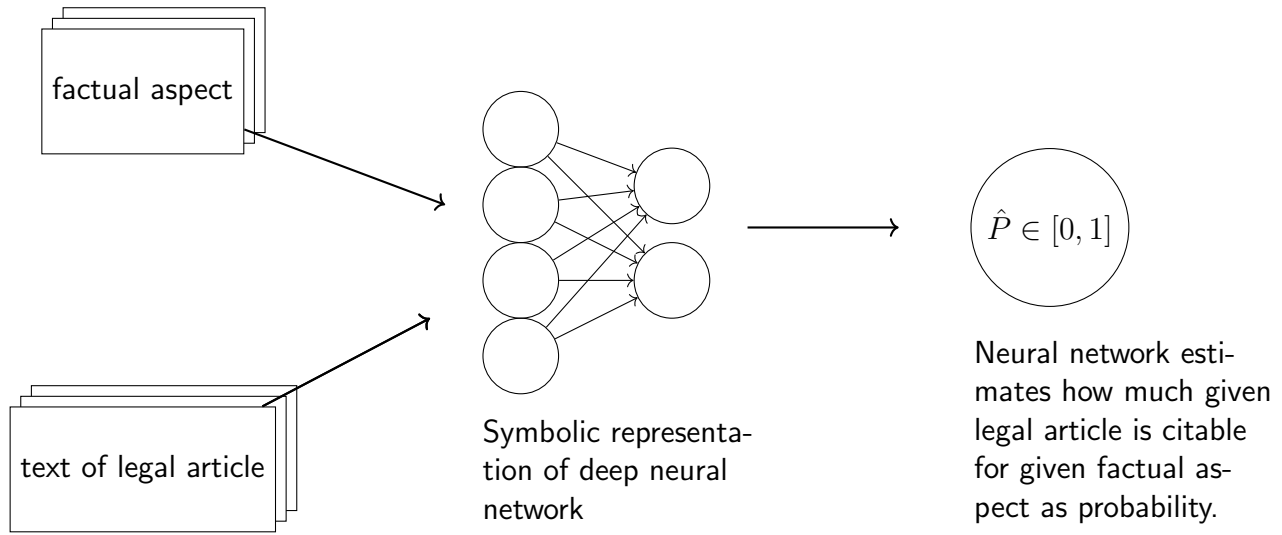


Figure 17: **Design of Training Framework of Deep Neural Network:** I designed a training framework of deep neural network by analogy with how member cites in WTO DSB as visualized in Figure 15 and Figure 16.

3. This set of directed edge weights W represents a set of conditional probability as conditional probability $P(v_j|v_i)$ how probably a source node v_i clarifies the meaning of the target node v_j compared to other source nodes as illustrated in Figure 4.

4 Empirical Findings

5 Conclusion

For given V , D defined in Figure 10 and Figure 11,
let $E = \{e \mid e \text{ is an english word or special charcter}\}$ and
 $n_{\text{factual}}, n_{\text{article}} \in \mathbb{N}$ represents *max token length* of factual aspect and legal article respectively .

Then define $T = \{t_d \mid t_d = (e_1, e_2, \dots, e_{n_{\text{factual}}}) \text{ s.t. } d \in D \text{ and } e_{i \leq n_{\text{factual}}} \in E\}$

where t represents a factual aspect of one of DS cases listed in Figure 8.

Also define $A = \{a_v \mid a_v = (e_1, e_2, \dots, e_{n_{\text{article}}}) \text{ s.t. } v \in V \text{ and } e_{i \leq n_{\text{article}}} \in E\}$

where t represents texts of one of a legal article listed in Figure 10.

Then define a deep neural network f with a set of parameters θ

$$f_{\theta} : T \times A \rightarrow [0, 1] \in \mathbb{R}$$

Figure 18: **Formal Definition of Input/Ouput of Deep Neural Network**

	Embedding Dimension = k					
This						Max Sequence Length = n
dispute						
concerns						
the						
:						
or						
issued						
thereafter						
[PAD]						

Figure 19: **Embedding Layer:** $n \times k$ matrix representation of factual aspect

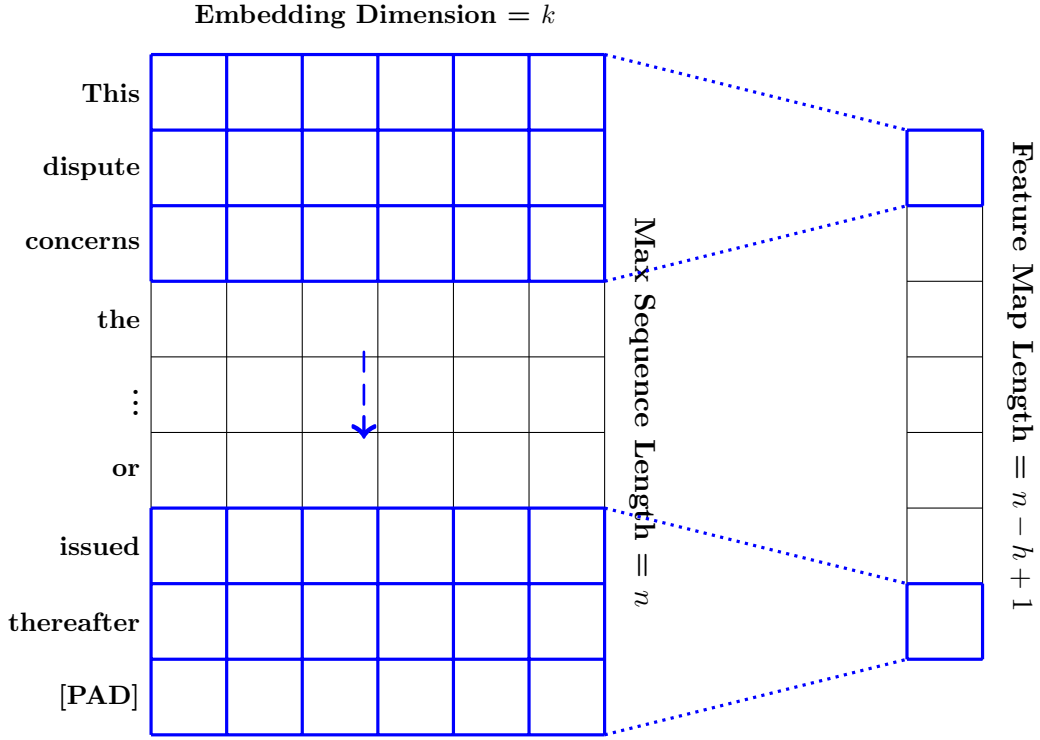


Figure 20: **Conv1D**: h -sized filter runs over the $n \times k$ embedding matrix and produces $(n - h + 1)$ size of feature map.

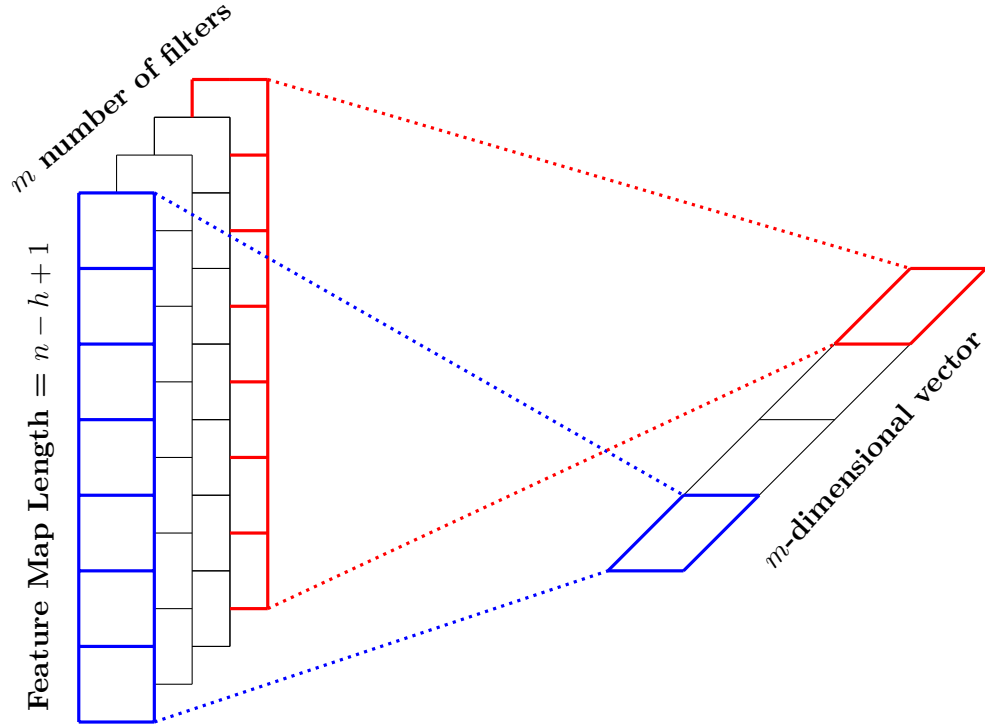
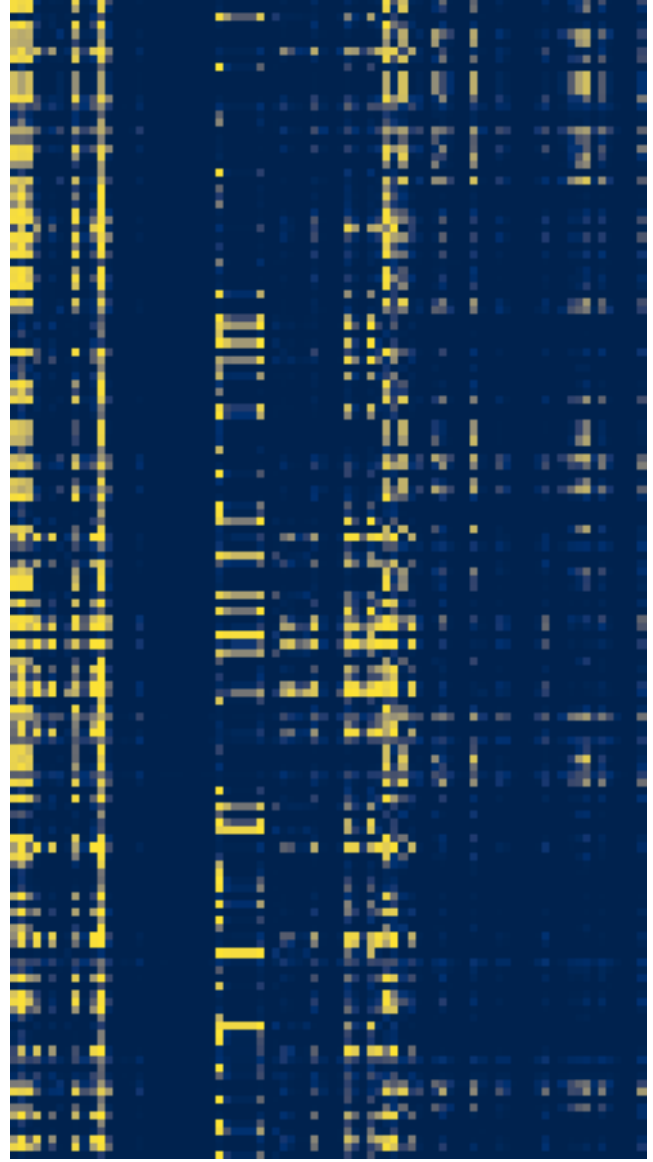


Figure 21: **MaxPool1D**: Filter out max value for all m number of feature map outputs from m different convolution filters. MaxPool1D produces m dimensional vector as an output for collection of those filtered max values.

	I	I:1	II	II:1	...
DS1	0	0.053	0.125	0.035	
DS2	0.040	0	0.04	0.054	
DS3	0.114	0.049	0	0.065	
DS4	0.032	0.065	0.065	0	
⋮					
DS505					
DS513					
DS518					
DS523					

(a) **Illustration of Prediction Matrix:** By defining row as list of DS case numbers and column as legal articles, we can create $|D| \times |V|$ matrix that includes $f_{\theta^*}(t_d, a_v)$ for each cell.



(b) Prediction

Figure 22: **Co-citation & Prediction**

References

- Charles, Zachary and Dimitris Papailiopoulos. 2017. “Stability and Generalization of Learning Algorithms that Converge to Global Optima.”.
- Dinh, Laurent, Razvan Pascanu, Samy Bengio and Yoshua Bengio. 2017. “Sharp Minima Can Generalize For Deep Nets.”.
- Fredebeul-Krein, Markus and Andreas Freytag. 1999. “The case for a more binding WTO agreement on regulatory principles in telecommunication markets.” *Telecommunications Policy* 23(9):625 – 644.
- Guo, Jiaxian, Sidi Lu, Han Cai, Weinan Zhang, Yong Yu and Jun Wang. 2017. “Long Text Generation via Adversarial Training with Leaked Information.”.
- Hufbauer, Gary and Sherry Stephenson. 2007. “Services Trade: Past Liberalization and Future Challenges.” *Journal of International Economic Law* 10(3):605–630.
- Huynh-Thu, Vân Anh, Alexandre Irrthum, Louis Wehenkel and Pierre Geurts. 2010. “Inferring Regulatory Networks from Expression Data Using Tree-Based Methods.” *PLOS ONE* 5(9):1–10.
- Johns, Leslie and Krzysztof J. Pelc. 2014. “Who Gets to Be In the Room? Manipulating Participation in WTO Disputes.” *International Organization* 68(3):663–699.
- Magdum, P. G. and Sheetal Rathi. 2021. A Survey on Deep Learning-Based Automatic Text Summarization Models. In *Advances in Artificial Intelligence and Data Engineering*, ed. Niranjan N. Chiplunkar and Takanori Fukao. Singapore: Springer Singapore pp. 377–392.
- Minaee, Shervin, Nal Kalchbrenner, Erik Cambria, Narjes Nikzad, Meysam Chenaghlu and Jianfeng Gao. 2020. “Deep Learning Based Text Classification: A Comprehensive Review.”.
- Neyshabur, Behnam, Srinadh Bhojanapalli, David McAllester and Nathan Srebro. 2017. “Exploring Generalization in Deep Learning.”.
- Palmeter, D. and P.C. Mavroidis. 2004. *Dispute Settlement in the World Trade Organization: Practice and Procedure*. Dispute Settlement in the World Trade Organization: Practice and Procedure Cambridge University Press.
- Pelc, Krzysztof J. 2014. “The Politics of Precedent in International Law: A Social Network Application.” *The American Political Science Review* 108(3):547–564.

- Shaffer, Gregory. 2004. *Power, governance, and the WTO: a comparative institutional approach*. Cambridge Studies in International Relations Cambridge University Press p. 130–160.
- Strezhnev, BuenoAnton. 2014. “Using Latent Space Models to Study International Legal Precedent: An Application to the WTO Dispute Settlement Body.” *American Political Science Association 2014 Annual Meeting* .
- World Trade Organization. 1999. *WTO Agreements Series*. Number no. 2 in “WTO Agreements Series” World Trade Organization.
- World Trade Organization. 2017. *A Handbook on the WTO Dispute Settlement System*. A WTO Secretariat publication Cambridge University Press.
- Zahavy, Tom, Bingyi Kang, Alex Sivak, Jiashi Feng, Huan Xu and Shie Mannor. 2017. “Ensemble Robustness and Generalization of Stochastic Deep Learning Algorithms.”.

Appendix A

A.1 Factual Aspect Example

Excerpt below is from the panel report for the *US - Offset Act (Byrd Amendment)*⁸ case.

II. FACTUAL ASPECTS

2.1 This dispute concerns the Continued Dumping and Subsidy Offset Act of 2000 (the “CDSOA” or the “Offset Act”), which was enacted on 28 October 2000 as part of the Agriculture, Rural Development, Food and Drug Administration and Related Agencies Appropriations Act, 2001.¹ The CDSOA amends Title VII of the Tariff Act of 1930 by adding a new section 754 entitled Continued Dumping and Subsidy Offset. Regulations prescribing administrative procedures under the Act were brought into effect on September 21, 2001.

2.2 The CDSOA provides that :

“ Duties assessed pursuant to a countervailing duty order, an anti-dumping duty order, or a finding under the Antidumping Act of 1921 shall be distributed on an annual basis under this section to the affected domestic producers for qualifying expenditures. Such distribution shall be known as “the continued dumping and subsidy offset”. ”

2.3 The term “affected domestic producers” means :

“ a manufacturer, producer, farmer, rancher, or worker representative (including associations of such persons) that –

(A) was a petitioner or interested party in support of the petition with respect to which an anti-dumping duty order, a finding under the Antidumping Act of 1921, or a countervailing duty order has been entered, and

(B) remains in operation.

⁸Panel Report, United States — Continued Dumping and Subsidy Offset Act of 2000, WTO Doc. WT/DS217/R (adopted Jan. 27, 2003).

Companies, business, or persons that have ceased the production of the product covered by the order or finding or who have been acquired by a company or business that is related to a company that opposed the investigation shall not be an affected domestic producer. ”

2.4 In turn, the term “qualifying expenditure” is defined by the CDSOA as “expenditure[s] incurred after the issuance of the anti-dumping duty finding or order or countervailing duty order in any of the following categories: “

- (A) Manufacturing facilities.
- (B) Equipment.
- (C) Research and development.
- (D) Personnel training.
- (E) Acquisition of technology.
- (F) Health care benefits to employees paid for by the employer.
- (G) Pension benefits to employees paid for by the employer.
- (H) Environmental equipment, training or technology.
- (I) Acquisition of raw materials and other inputs.
- (J) Working capital or other funds needed to maintain production.” ”

2.5 The CDSOA provides that the Commissioner of Customs shall establish in the Treasury of the United States a special account with respect to each order or finding⁸ and deposit into such account all the duties assessed under that Order.⁹ The Commissioner of Customs shall distribute all funds (including all interest earned on the funds) from the assessed duties received in the preceding fiscal year to affected domestic producers based on a certification by the affected domestic producer that he is eligible to receive the distribution and desires to receive a distribution for qualifying expenditures incurred since the issuance of the order or finding.¹⁰ Funds deposited in each special account during each fiscal year are to be distributed no later than 60 days after the beginning of the following fiscal year.¹¹ The CDSOA and regulations prescribe that (1) if the total amount of the certified net claims filed by affected domestic producers does not exceed the amount of the offset available, the certified net claim for each affected domestic producer will be paid in full, and (2) if the certified net claims exceed the amount available, the offset will be made

on a pro rata basis based on each affected domestic producer's total certified claim.

2.6 Special accounts are to be terminated after “(A) the order or finding with respect to which the account was established has terminated; (B) all entries relating to the order or finding are liquidated and duties assessed collected; (C) the Commissioner has provided notice and a final opportunity to obtain distribution pursuant to subsection (c); and (D) 90 days has elapsed from the date of the notice described in subparagraph (C).” All amounts that remain unclaimed in the Account are to be permanently deposited into the general fund in the US Treasury.¹²

2.7 The CDSOA applies with respect to all anti-dumping and countervailing duty assessments made on or after 1 October 2000¹³ pursuant to an anti-dumping order or a countervailing order or a finding under the Antidumping Act of 1921 in effect on 1 January 1999 or issued thereafter. [END]

A.2 Collected Cited Articles for 143 WTO DSB Cases

DS refers to *Dispute Settlement* and this notation is officially adopted by WTO DSB.

WTO DSB identifies each dispute with a unique number for each case such as DS2 and DS18.

Case Number	Cited Articles (GATT 1994)
DS 2	I, III, XXII:1
DS 18	XI, XIII
DS 22	VI:3, VI:6
DS 31	III, XI
DS 34	XI, XIII, XXIV
DS 46	XVI
DS 56	II, VII, VIII, X
DS 58	I, XI, XIII, XX
DS 60	VI
DS 62	II
DS 67	II, XXIII, XXIII:1

DS 68	II, XXII:1, XXIII:1
DS 69	II, III, X, XIII, XXVIII
DS 75	III:2
DS 76	XI
DS 87	III:2
DS 90	XI:1, XIII, XVIII:11
DS 98	XIX
DS 103	X, XI, XIII
DS 108	III:4, XVI
DS 121	XIX
DS 122	VI
DS 135	III, XI, XXIII, XXIII:1(b)
DS 136	III:4, VI
DS 139	I:1, III:4, XXIV
DS 141	I, VI
DS 146	III, XI
DS 152	I, II, III, VIII, XI
DS 155	III:2, X:3(a), XI:1
DS 161	II, III, X, XI, XVII
DS 162	III, III:4, VI, XI
DS 165	I, II, VIII, XI
DS 166	I, XIX
DS 174	I, III:4
DS 175	III, III:4, XI, XI:1
DS 177	I, II, XIX
DS 184	VI, X
DS 202	I, XIII, XIX
DS 207	II, XIX:1
DS 212	VI:3
DS 217	VI:2, VI:3, X:3, XXIII:1

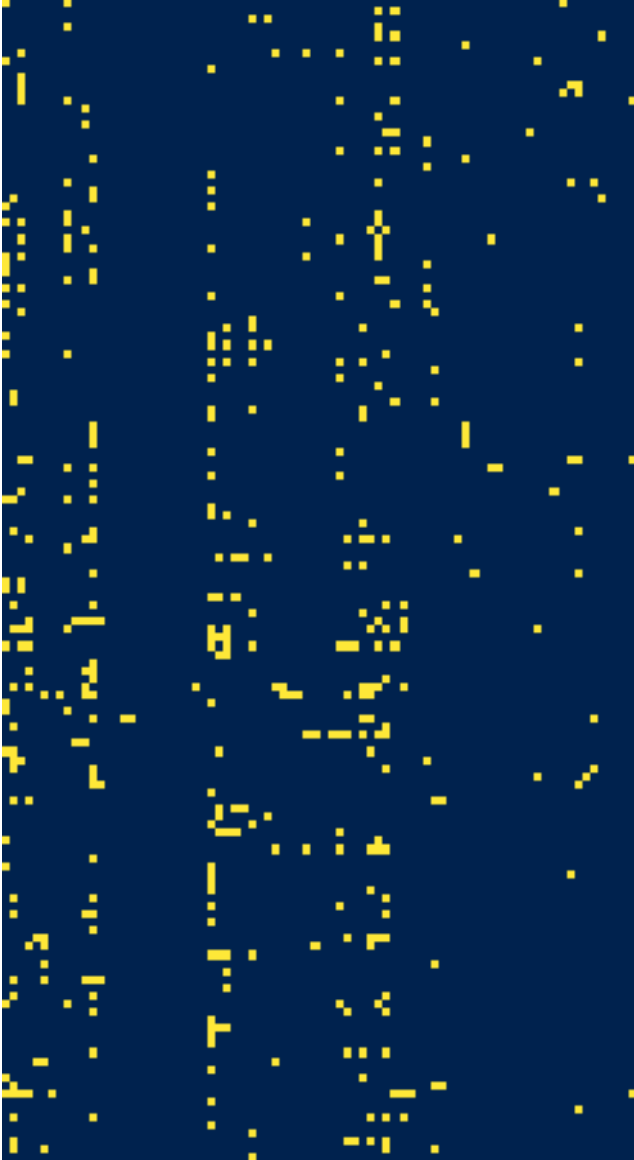
DS 219	I, VI
DS 221	VI, VI:2, VI:3, VI:6
DS 231	I, III, XI:1
DS 234	VI, VI:2, VI:3, X, X:3, XXIII:1
DS 238	XIX:1
DS 244	VI, X
DS 245	XI
DS 246	I:1
DS 248	I:1, XIII, XIX:1
DS 257	VI, VI:3, X:3
DS 264	VI, X:3
DS 265	III:4, XVI
DS 266	III:4, XVI
DS 267	III:4, XVI
DS 268	VI, X
DS 269	II, II:1, XXIII, XXIII:1, XXVIII
DS 276	III, III:4, XVII, XVII:1
DS 282	VI, X
DS 283	III:4
DS 286	II, XXII
DS 290	I, I:1, III, III:4
DS 294	VI
DS 295	VI, VI:2
DS 296	VI:3, X:3
DS 301	I:1, III:4, XXIII:1
DS 302	II:1, III:2, III:4, X:1, X:3, X:3(a), XI:1, XV
DS 308	III
DS 312	VI:1, VI:2(a), VI:2(b), VI:6
DS 315	X:1, X:3
DS 316	III:4, XVI:1, XXIII:1

DS 320	I, II
DS 321	I, II
DS 322	VI, VI:1, VI:2(a)
DS 332	I:1, III:4, XI:1, XIII:1
DS 336	VI:3, X:3
DS 339	II:1, III:1, III:2, III:4, III:5, XI, XIII:1
DS 343	I:1, II, II:1, III, VI, VI:2, X:3(a), XI:1, XIII:1, XX
DS 344	VI, VI:1, VI:2
DS 345	I, II, II:1, VI, VI:2, VI:3, X, X:1, X:2, XI, XIII
DS 350	VI:1, VI:2
DS 353	III:4
DS 360	II:1, III:2, III:4
DS 363	III:4, XI:1
DS 366	I:1, II:1, III:2, V:6, VII, VII:1, X:3, X:3(a), XI, XIII:1
DS 371	II:1(b), II:3, III:2, III:4, VII:1, VII:2, VII:5, X:1, X:3, X:3(a)
DS 379	I, VI
DS 381	I, III
DS 384	III:4, IX, IX:2, X:3, X:3(a), XXIII:1(b)
DS 392	I:1, XI:1
DS 394	VIII, VIII:1, VIII:4, X, X:1, X:3, XI, XI:1
DS 396	III:1, III:2
DS 397	I, I:1, VI:1, X:3(a)
DS 399	I:1, II, XIX
DS 400	I:1, III:4, XI:1, XXIII:1(b)
DS 406	III:4, XX, XXIII:1(a)
DS 412	III:4, III:5, XXIII:1
DS 414	VI
DS 415	I:1, II:1, XIX:1, XIX:2
DS 422	VI:1, VI:2(a), VI:2(b)
DS 425	VI:1, VI:6

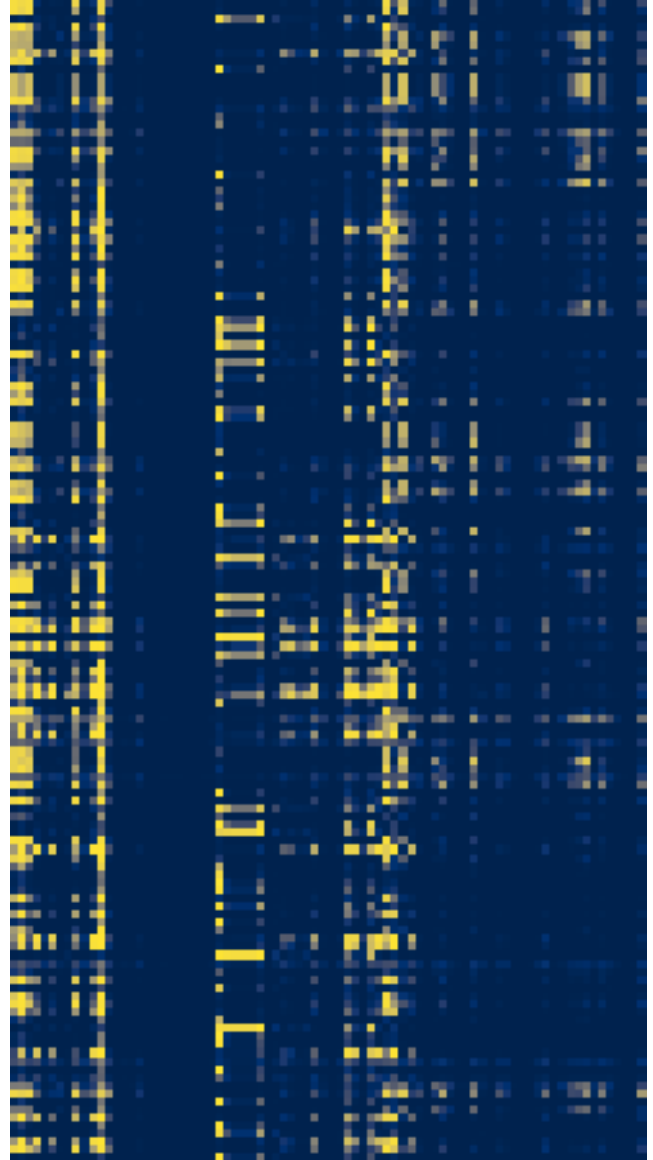
DS 427	VI, VI:3
DS 429	VI:1, VI:2, VI:2(a), X
DS 430	I, XI
DS 431	VII, VIII, X, X:3(a), XI, XI:1
DS 435	III:4
DS 436	I, VI
DS 437	VI, XXIII
DS 440	VI
DS 442	VI, X:3(a)
DS 447	I:1, III:4, XI:1
DS 449	VI, X
DS 453	I:1, III:2, III:4, XI:1
DS 454	VI
DS 456	III:4
DS 457	II:1(a), II:1(b), X:1, X:3(a), XI, XI:1
DS 461	II:1, II:1(b), VIII:1, X:3(a)
DS 464	VI, VI:1, VI:2, VI:3
DS 468	II:1(b), XIX:1
DS 471	VI:2
DS 472	I:1, II:1(b), III:2, III:4, III:5
DS 473	VI:2
DS 475	I:1, III:4, XI:1
DS 476	I, III, X, XI
DS 477	III:4, X:1, XI:1
DS 479	VI
DS 480	VI, VI:1, VI:2
DS 482	VI
DS 483	VI
DS 484	III:4, X:1, X:3, XI:1
DS 485	II:1(a), II:1(b), VII

DS 486	VI
DS 488	I, X:3
DS 490	I:1, XIX:1, XIX:2
DS 492	I, I:1, II, II:1, II:2, XIII, XIII:1, XIII:2, XXVIII
DS 493	VI
DS 495	XXIII:1
DS 499	I:1, III:4, X:3(a), XI:1, XIII:1
DS 504	VI
DS 505	VI:3
DS 513	I:1, X:1, X:2, X:3(a), XI:1
DS 518	I:1, II:1(b), XI:1, XIX:1
DS 523	VI:3

A.3 Technical Details



(a) Co-citation



(b) Prediction

Figure 23: **Co-citation & Prediction**