

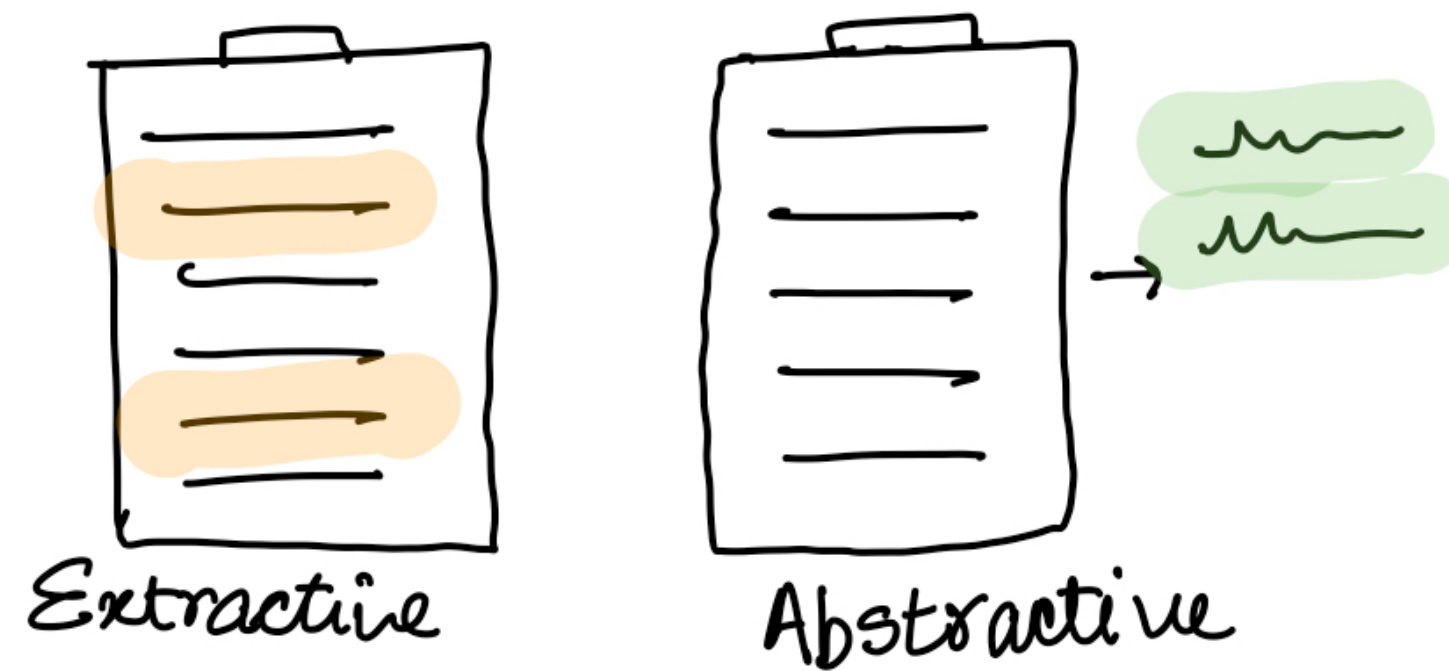
Abstractive Summarization

Wikipedia



Introduction

Introduction



Abstractive summarization = summaries with **novel** phrases and words.

Best artificial abstractive summaries achieved with neural networks,
but with a huge gap to human performance!

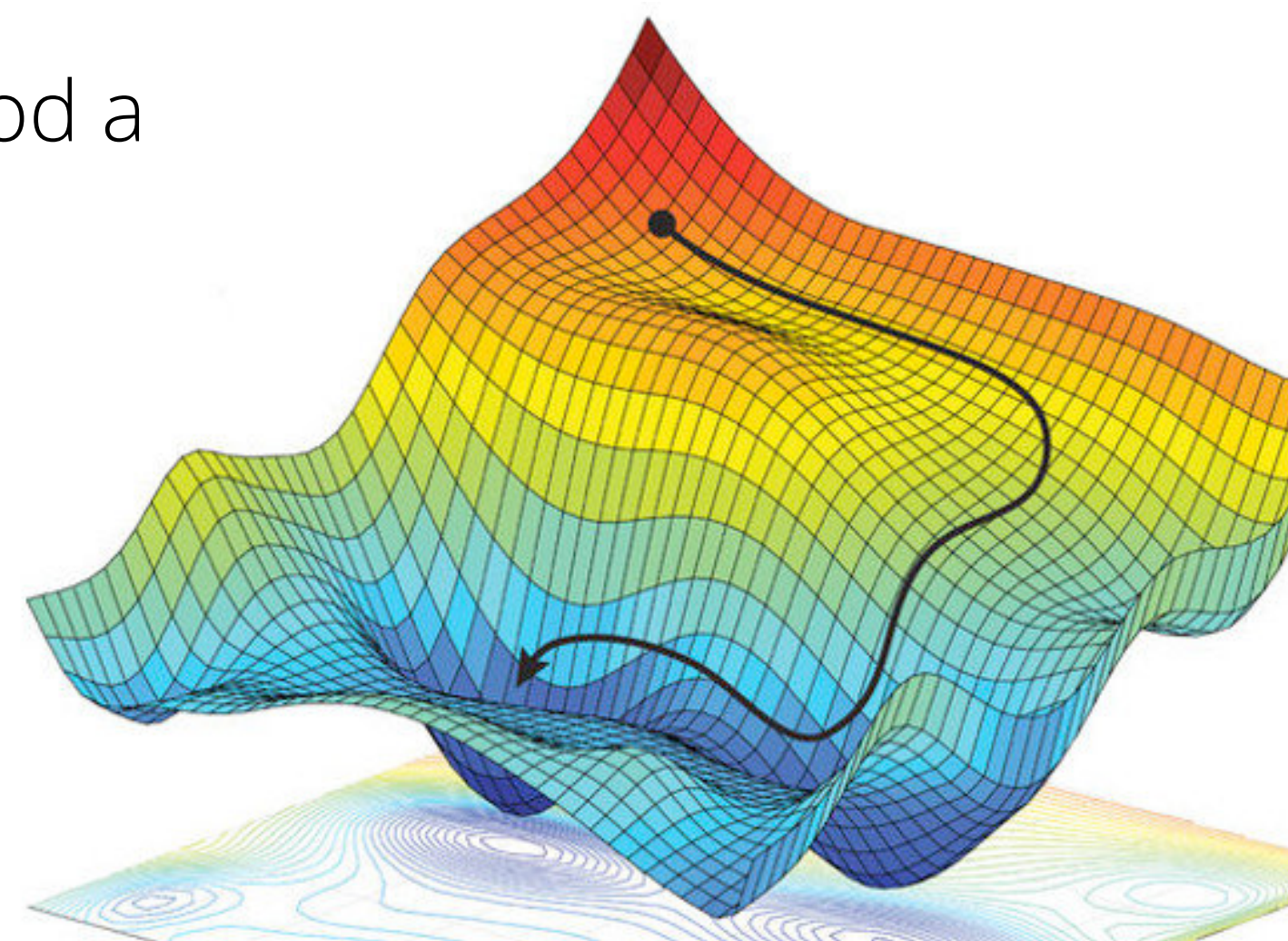
Choosing a loss function

Neural networks need loss function to train.

Next word prediction is simple but has some problems.

Human experts would give best feedback on how good a summary is, but not usable as loss.

- ▶ Using another model for evaluating summaries.
- ▶ Model feedback has potential as a metric!

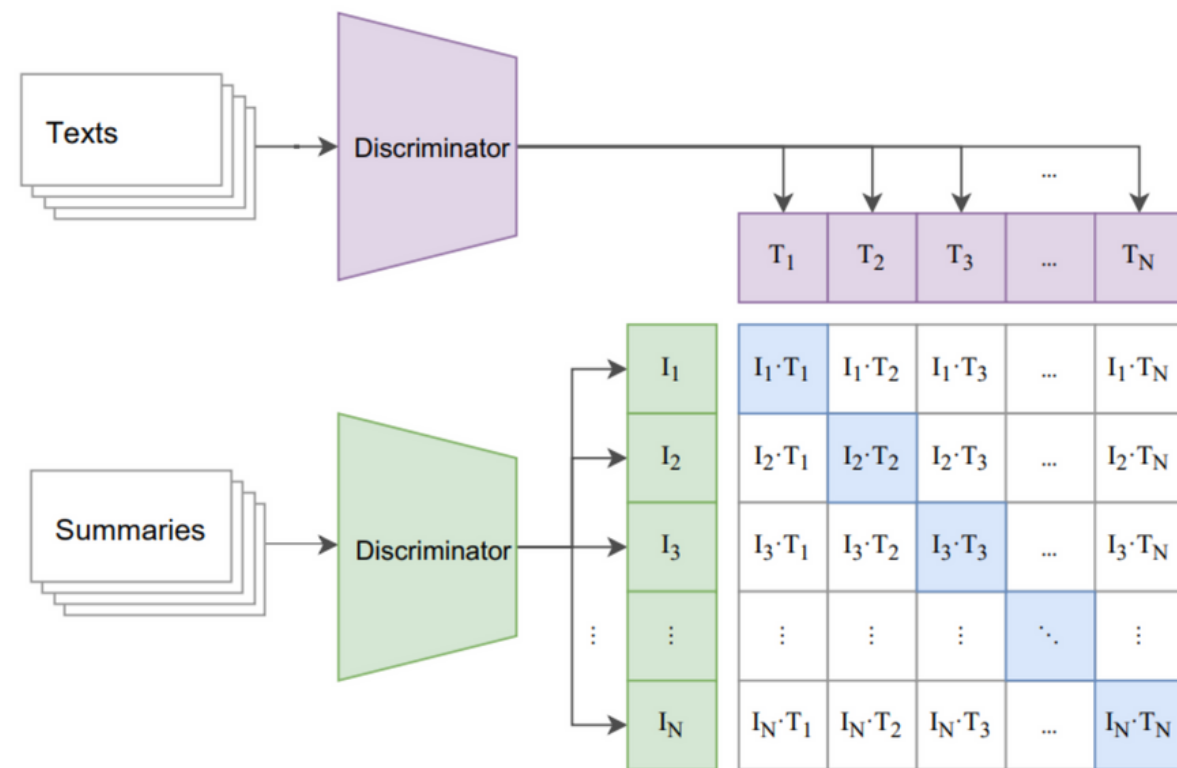


Architecture and Pipeline

Overview approach

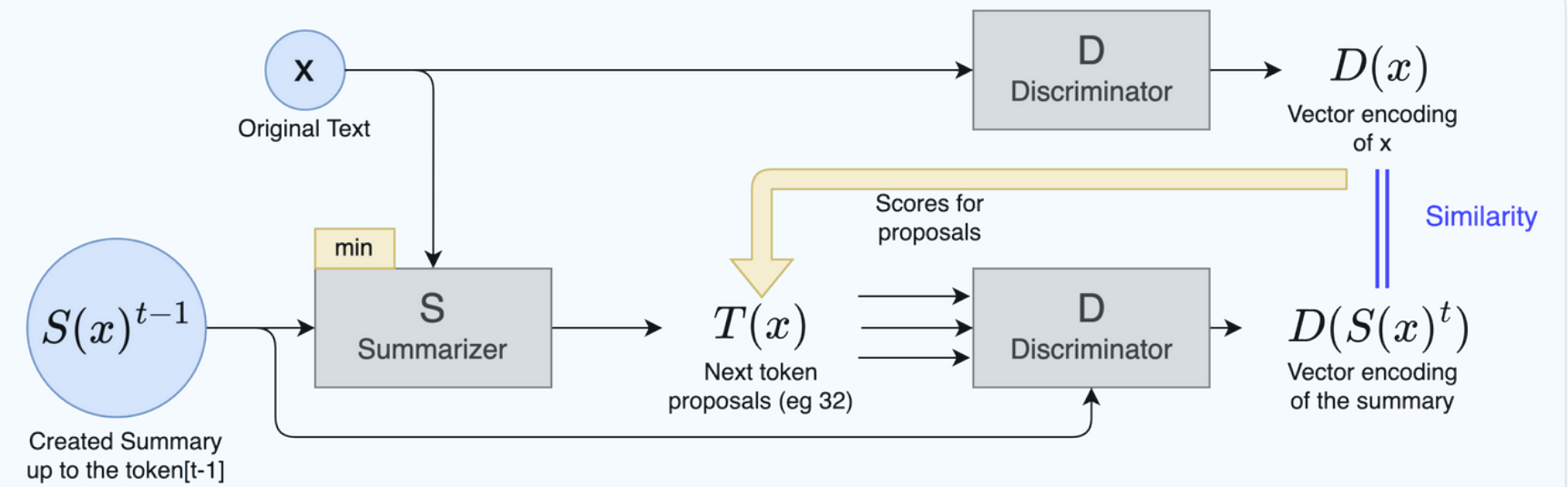
Step 0: Contrastive-training of discriminator

(1) Contrastive pre-training

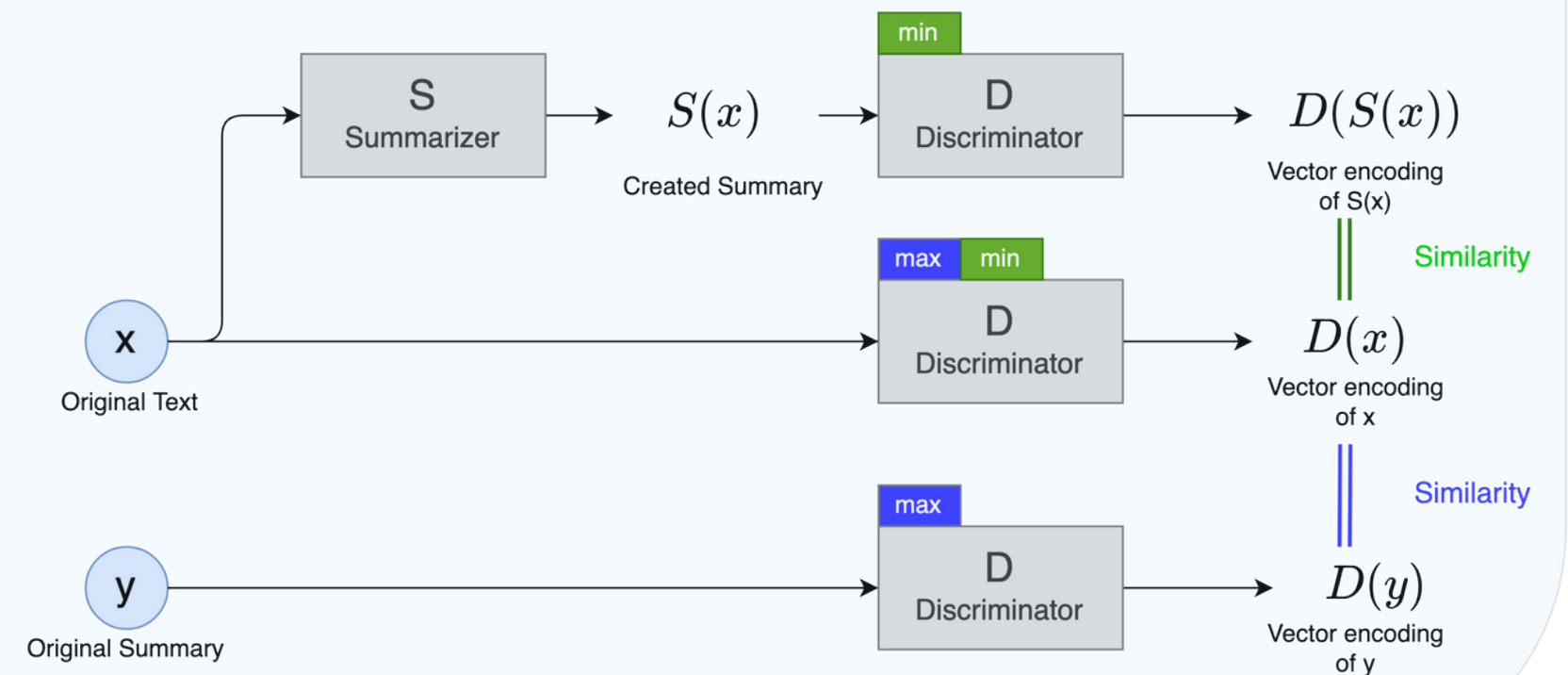


GAN Loop

Step 1: Training the summary generator



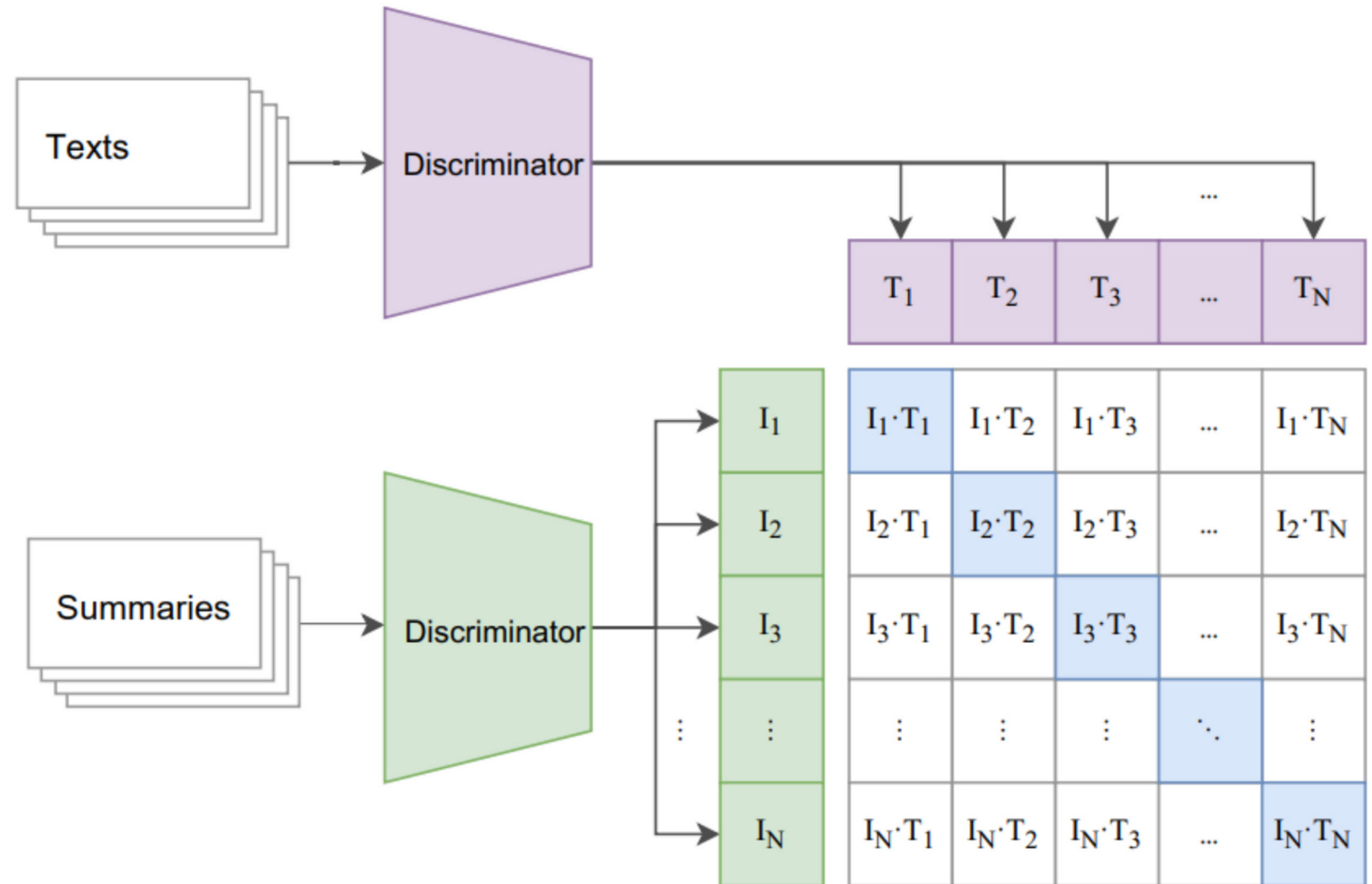
Step 2: Training the discriminator



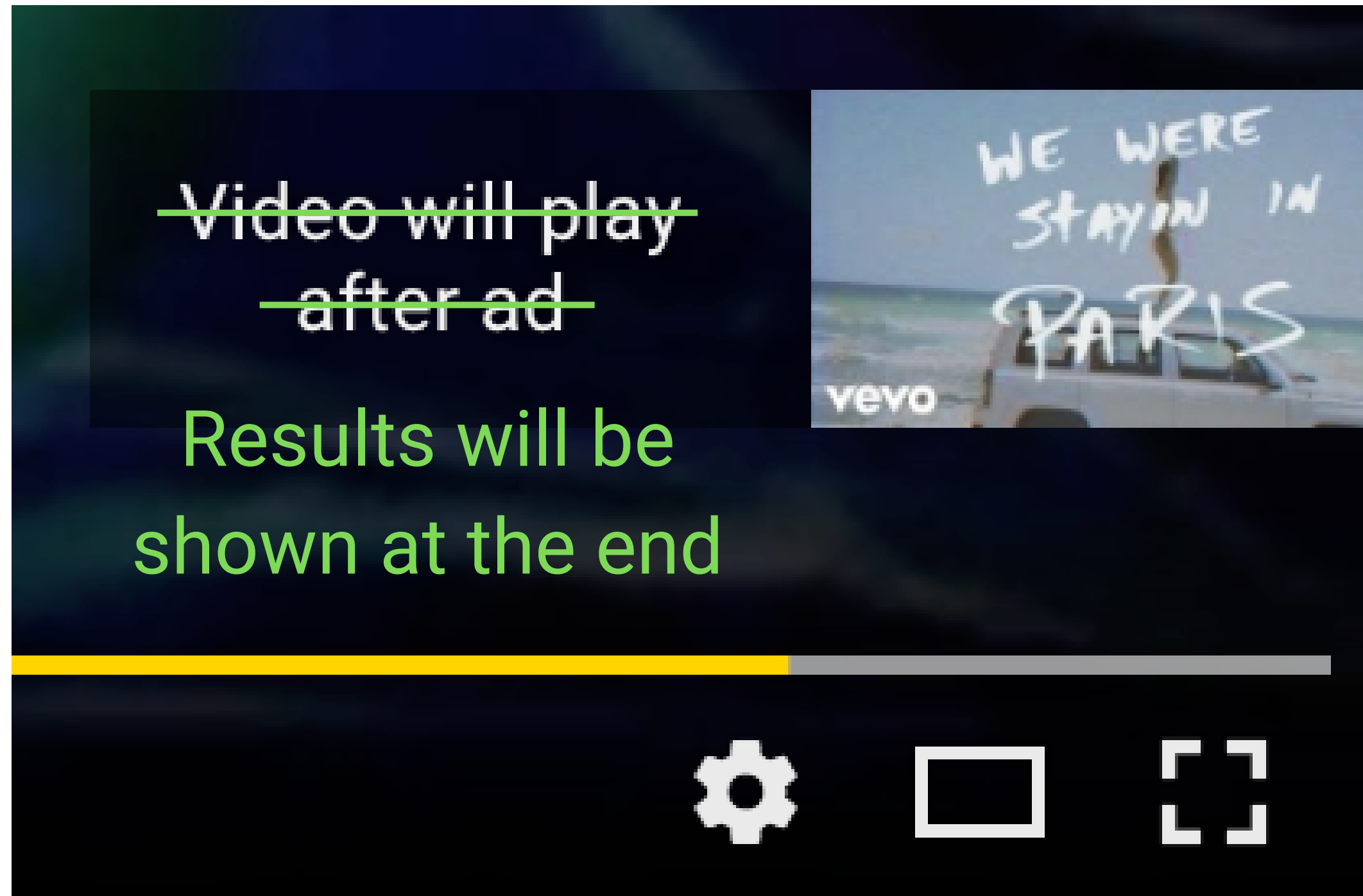
Step 0: Contrastive-training of discriminator

Idea behind:

- Encourage matches.
- Discourage mismatches.



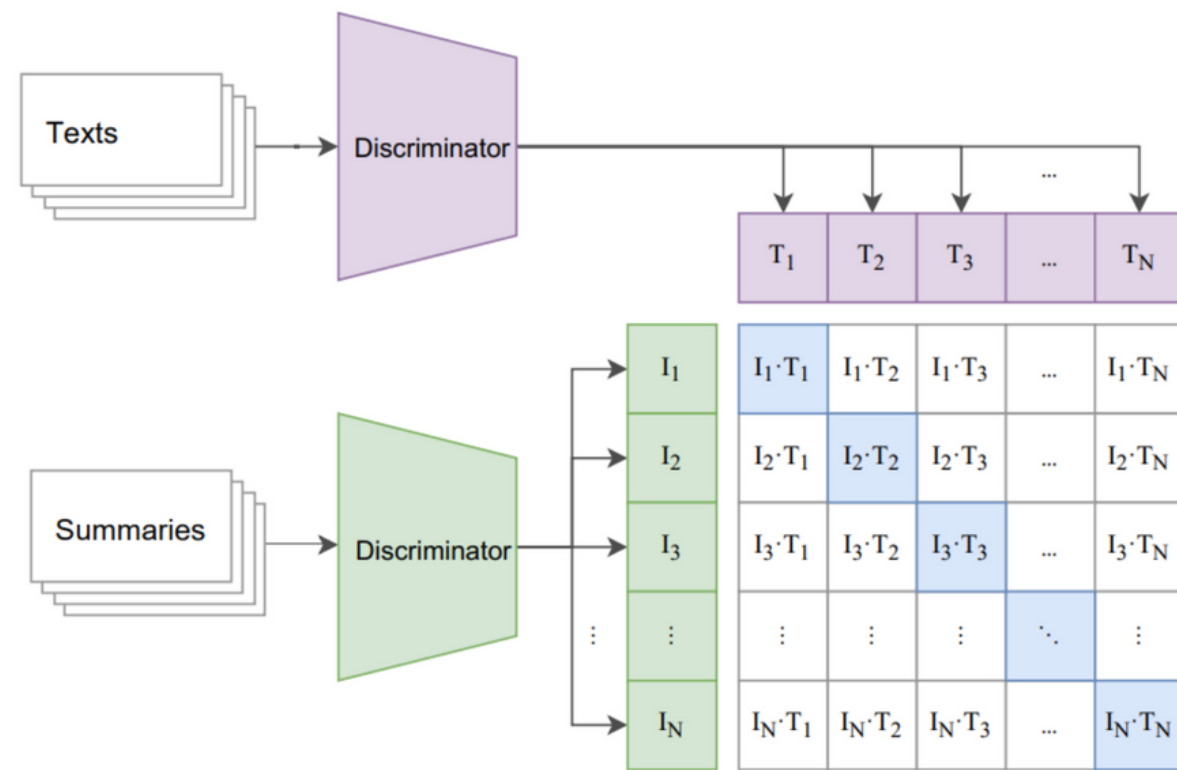
First successes: Discriminator as metric



Review approach

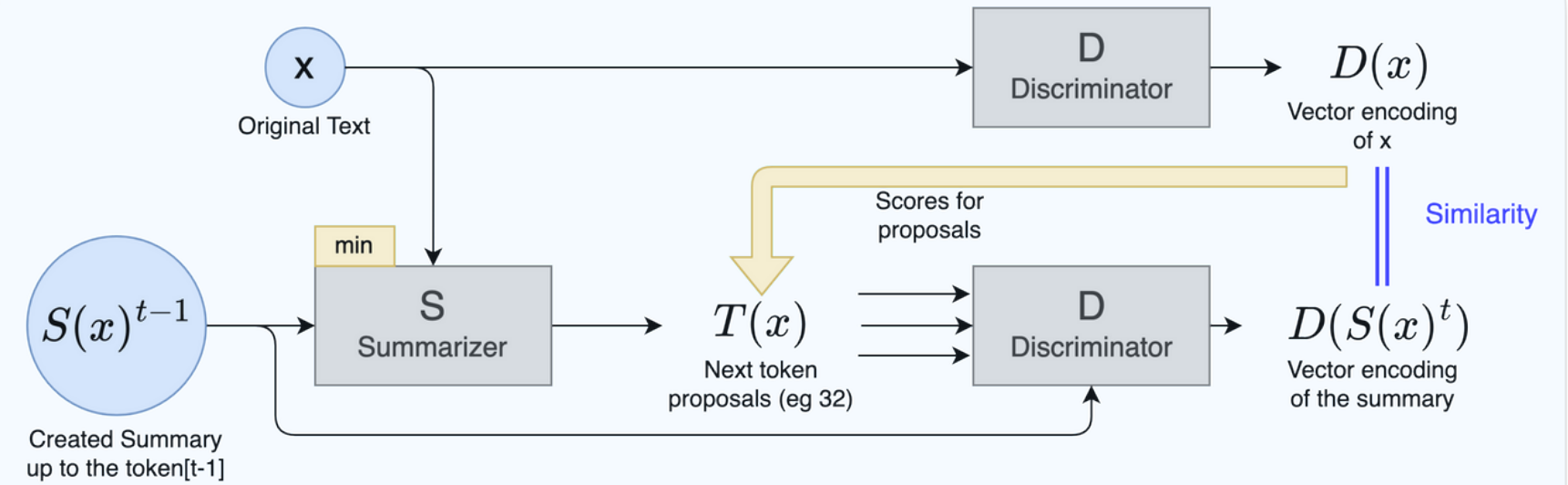
Step 0: Contrastive-training of discriminator

(1) Contrastive pre-training

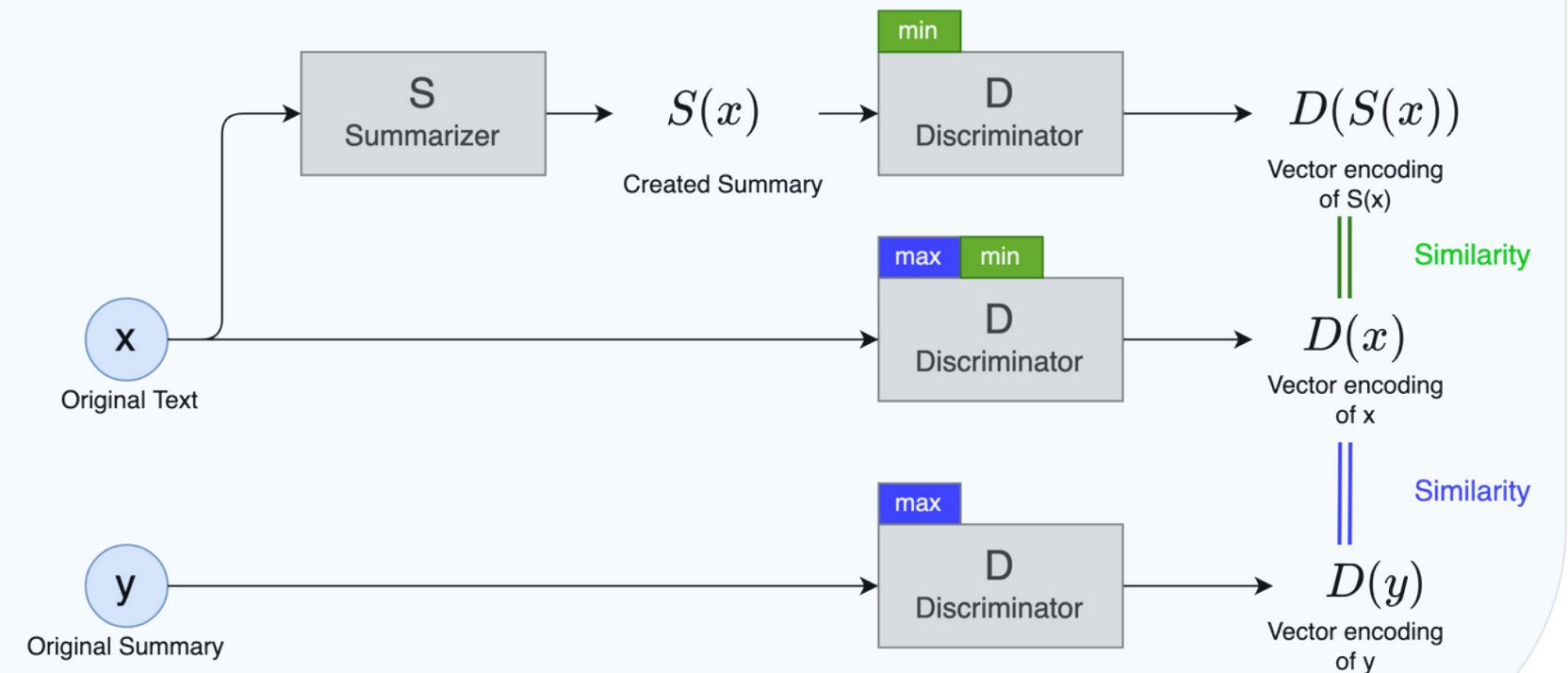


GAN Loop

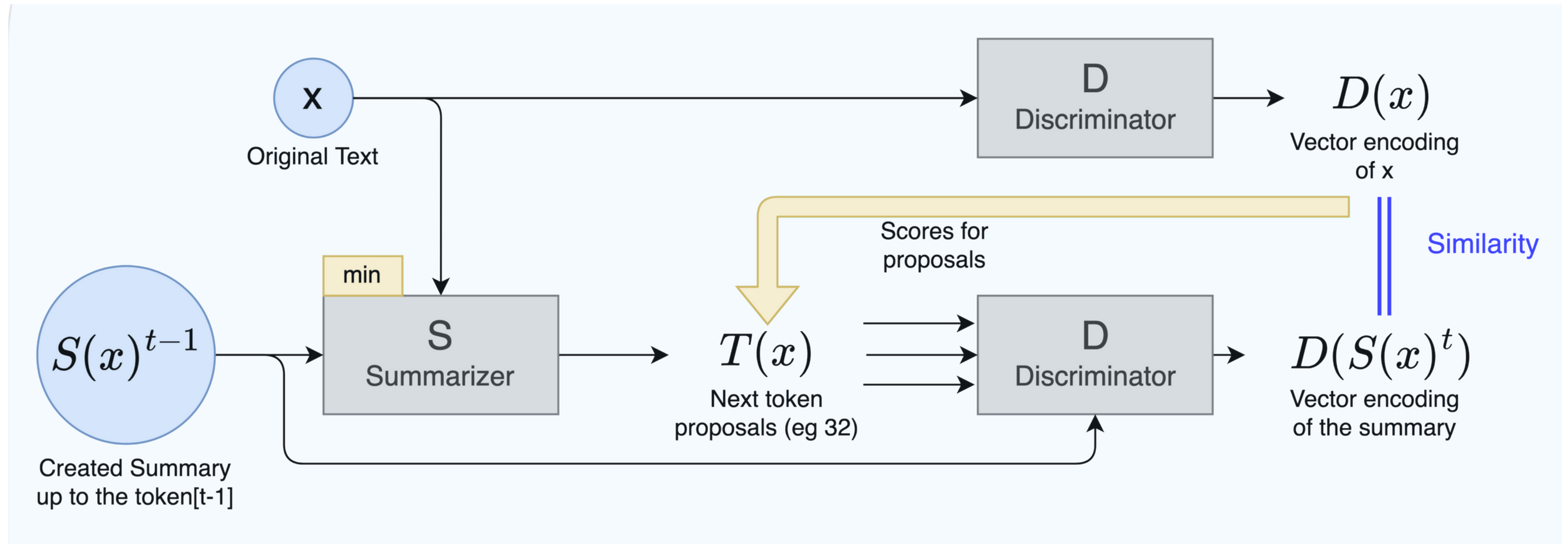
Step 1: Training the summary generator



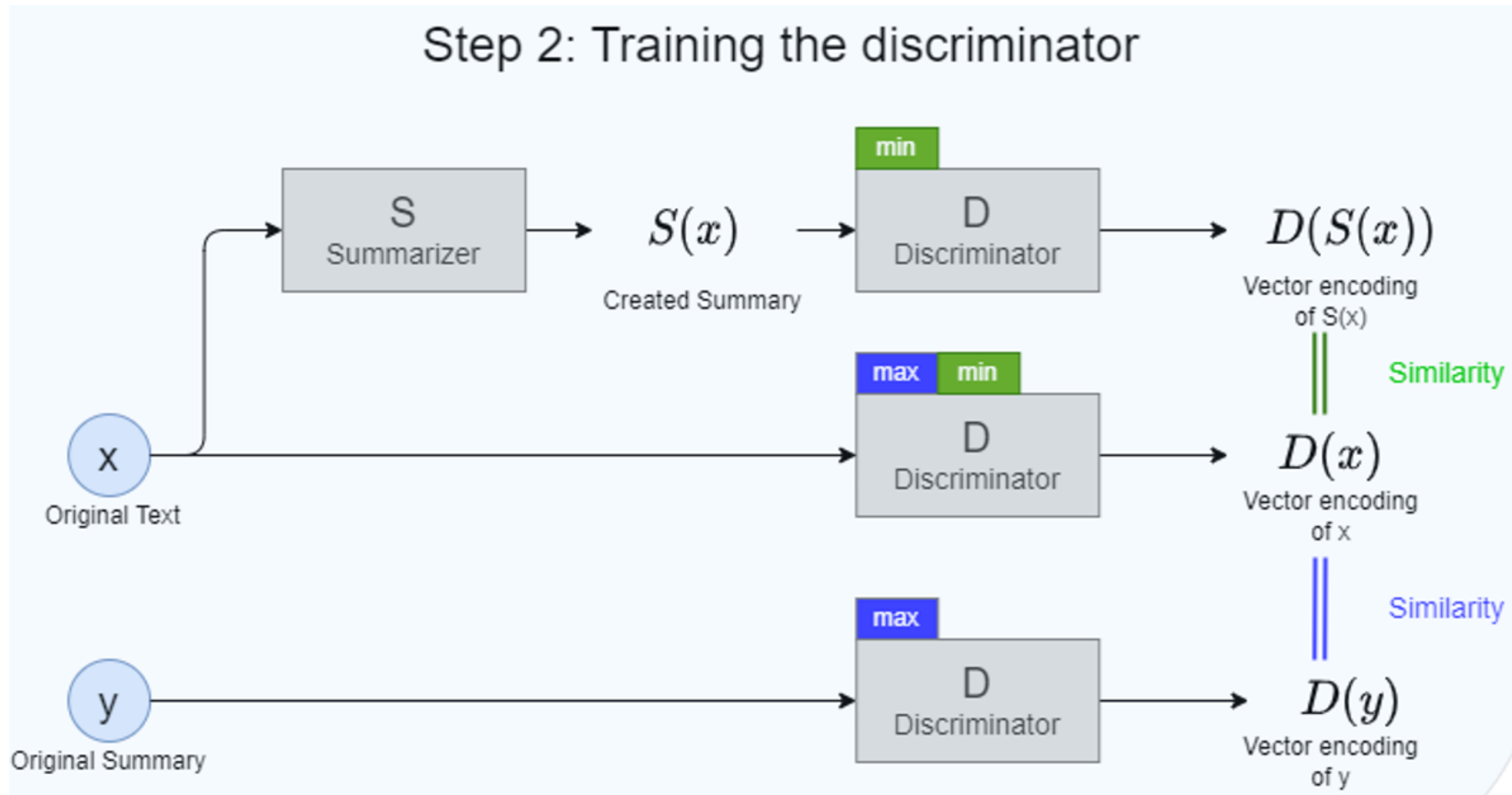
Step 2: Training the discriminator



Step 1: Training of the summary generator



Step 2: Training of the discriminator



Datasets

Datasets (I)



CNN & DM Dataset



Hugging Face



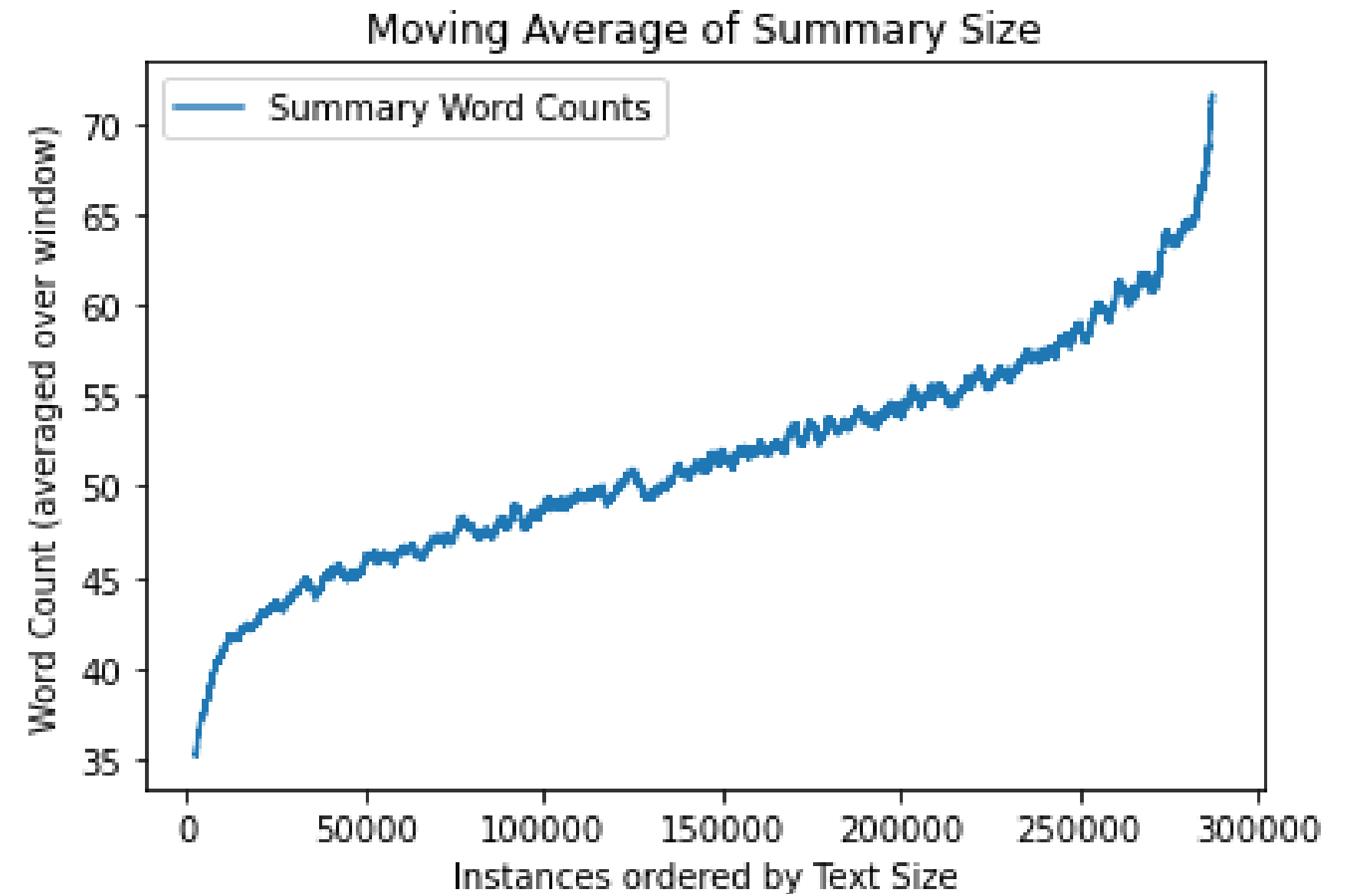
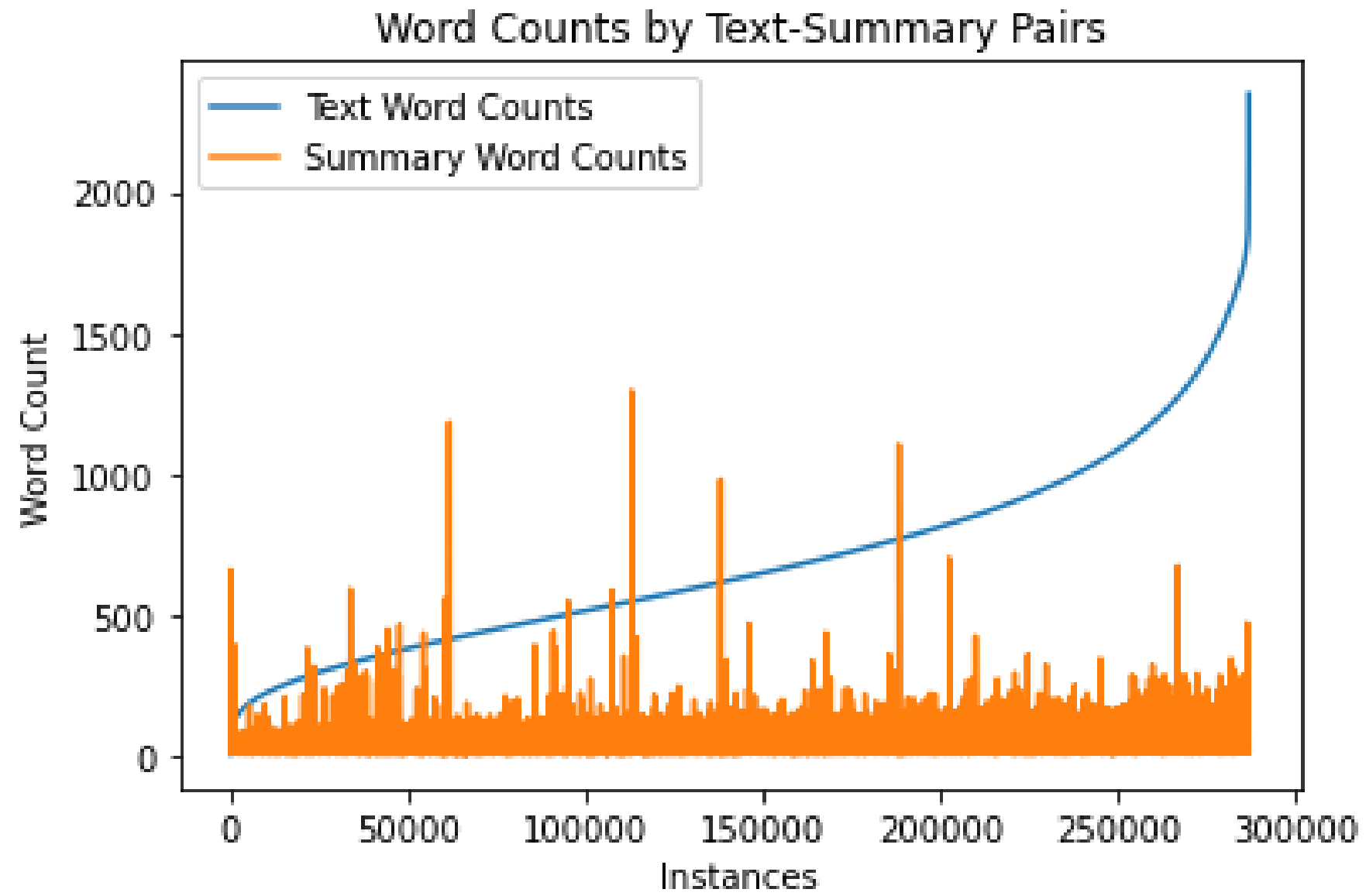
Wikipedia Summarization Dataset



Datasets (II)



CNN & DM Dataset



Paul Stäckel

From Wikipedia, the free encyclopedia

Paul Gustav Samuel Stäckel (20 August 1862, [Berlin](#) – 12 December 1919, [Heidelberg](#)) was a German [mathematician](#), active in the areas of [differential geometry](#), [number theory](#), and [non-Euclidean geometry](#). In the area of [prime number theory](#), he used the term *twin prime* (in its [German](#) form, "Primzahlzwilling") for the first time.^{[1][2]}

After passing his *Abitur* in 1880 he studied mathematics and physics at the [University of Berlin](#), but also listened to lectures on philosophy, psychology, education, and history. A year later he qualified for teaching in higher education and then taught at *Gymnasien* in Berlin. In 1885 he wrote his doctoral dissertation under [Leopold Kronecker](#) and [Karl Weierstraß](#). In 1891 he completed his *Habilitation* at the [University of Halle](#). Later he worked as a professor at the [University of Königsberg](#) (*außerordentlicher Professor* from 1895 to 1897), the [University of Kiel](#) (*ordentlicher Professor*, 1897 to 1905), [University of Hannover](#) (1905 to 1908), the [Karlsruhe Institute of Technology](#) (1908 to 1913), and the [University of Heidelberg](#) (1913 to 1919).

Stäckel worked on both mathematics and the history of mathematics. He edited the letters exchanged between [Carl Friedrich Gauss](#) and [Wolfgang Bolyai](#), made contributions to editions of the collected works of Euler and Gauss (for whose works he wrote *Gauss als Geometer*), and edited the *Geometrischen Untersuchungen* by Wolfgang and Johann Bolyai (published in 1913). Additionally he translated works of [Jacob Bernoulli](#), [Johann Bernoulli](#), [Augustin Louis Cauchy](#), [Leonhard Euler](#), [Joseph-Louis Lagrange](#), [Adrien-Marie Legendre](#), [Carl Gustav Jacobi](#) from French and Latin into German for the series *Ostwalds Klassiker der exakten Wissenschaften*.

In 1904 he was an invited speaker at the [International Congress of Mathematicians](#) in Heidelberg.^[3] In 1905 he was the president of the [Deutsche Mathematiker-Vereinigung](#). His doctoral students include [Paul Riebesell](#).



Works [\[edit\]](#)

- *Über die Bewegung eines Punktes auf einer Fläche*^[?], 1885, Dissertation
- *Die Integration der Hamilton-Jacobischen Differentialgleichung mittelst Separation der Variablen*, 1891, Habilitation
- *Elementare Dynamik der mathematischen Wissenschaft* in: Encyclopädie der mathematischen Wissenschaft IV,1 (1908)
- several papers in the *Mathematischen Annalen* (from 1890 to 1909)
- numerous papers in "Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse" from 1896 to 1917.
- with [Friedrich Engel](#): *Theorie der Parallellinien von Euklid bis auf Gauss*^[?], 1895

Paul Stäckel

Summary

From Wikipedia, the free encyclopedia

Paul Gustav Samuel Stäckel (20 August 1862, [Berlin](#) – 12 December 1919, [Heidelberg](#)) was a German [mathematician](#), active in the areas of [differential geometry](#), [number theory](#), and [non-Euclidean geometry](#). In the area of [prime number theory](#), he used the term *twin prime* (in its [German](#) form, "Primzahlzwilling") for the first time.^{[1][2]}

After passing his *Abitur* in 1880 he studied mathematics and physics at the [University of Berlin](#), but also listened to lectures on philosophy, psychology, education, and history. A year later he qualified for teaching in higher education and then taught at [Gymnasien](#) in Berlin. In 1885 he wrote his doctoral dissertation under [Leopold Kronecker](#) and [Karl Weierstraß](#). In 1891 he completed his *Habilitation* at the [University of Halle](#). Later he worked as a professor at the [University of Königsberg](#) (*außerordentlicher Professor* from 1895 to 1897), the [University of Kiel](#) (*ordentlicher Professor*, 1897 to 1905), [University of Hannover](#) (1905 to 1908), the [Karlsruhe Institute of Technology](#) (1908 to 1913), and the [University of Heidelberg](#) (1913 to 1919).

Stäckel worked on both mathematics and the history of mathematics. He edited the letters exchanged between [Carl Friedrich Gauss](#) and [Wolfgang Bolyai](#), made contributions to editions of the collected works of Euler and Gauss (for whose works he wrote *Gauss als Geometer*), and edited the *Geometrischen Untersuchungen* by Wolfgang and Johann Bolyai (published in 1913). Additionally he translated works of [Jacob Bernoulli](#), [Johann Bernoulli](#), [Augustin Louis Cauchy](#), [Leonhard Euler](#), [Joseph-Louis Lagrange](#), [Adrien-Marie Legendre](#), [Carl Gustav Jacobi](#) from French and Latin into German for the series *Ostwalds Klassiker der exakten Wissenschaften*.

In 1904 he was an invited speaker at the [International Congress of Mathematicians](#) in Heidelberg.^[3] In 1905 he was the president of the [Deutsche Mathematiker-Vereinigung](#). His doctoral students include [Paul Riebesell](#).



Paul Stäckel.

Works [[edit](#)]

- *Über die Bewegung eines Punktes auf einer Fläche*^[4], 1885, Dissertation
- *Die Integration der Hamilton-Jacobischen Differentialgleichung mittelst Separation der Variablen*, 1891, Habilitation
- *Elementare Dynamik der mathematischen Wissenschaft* in: *Encyclopädie der mathematischen Wissenschaft* IV,1 (1908)
- several papers in the *Mathematischen Annalen* (from 1890 to 1909)
- numerous papers in "Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse" from 1896 to 1917.
- with [Friedrich Engel](#): *Theorie der Parallellinien von Euklid bis auf Gauss*^[5], 1895

Paul Stäckel

Summary

From Wikipedia, the free encyclopedia

Paul Gustav Samuel Stäckel (20 August 1862, [Berlin](#) – 12 December 1919, [Heidelberg](#)) was a German [mathematician](#), active in the areas of [differential geometry](#), [number theory](#), and [non-Euclidean geometry](#). In the area of [prime number theory](#), he used the term *twin prime* (in its [German](#) form, "Primzahlzwillig") for the first time.^{[1][2]}

After passing his *Abitur* in 1880 he studied mathematics and physics at the [University of Berlin](#), but also listened to lectures on philosophy, psychology, education, and history. A year later he qualified for teaching in higher education and then taught at [Gymnasien](#) in Berlin. In 1885 he wrote his doctoral dissertation under [Leopold Kronecker](#) and [Karl Weierstraß](#). In 1891 he completed his *Habilitation* at the [University of Halle](#). Later he worked as a professor at the [University of Königsberg](#) (*außerordentlicher Professor* from 1895 to 1897), the [University of Kiel](#) (*ordentlicher Professor*, 1897 to 1905), [University of Hannover](#) (1905 to 1908), the [Karlsruhe Institute of Technology](#) (1908 to 1913), and the [University of Heidelberg](#) (1913 to 1919).

Stäckel worked on both mathematics and the history of mathematics. He edited the letters exchanged between [Carl Friedrich Gauss](#) and [Wolfgang Bolyai](#), made contributions to editions of the collected works of Euler and Gauss (for whose works he wrote *Gauss als Geometer*), and edited the *Geometrischen Untersuchungen* by Wolfgang and Johann Bolyai (published in 1913). Additionally he translated works of [Jacob Bernoulli](#), [Johann Bernoulli](#), [Augustin Louis Cauchy](#), [Leonhard Euler](#), [Joseph-Louis Lagrange](#), [Adrien-Marie Legendre](#), [Carl Gustav Jacobi](#) from French and Latin into German for the series *Ostwalds Klassiker der exakten Wissenschaften*.

In 1904 he was an invited speaker at the [International Congress of Mathematicians](#) in Heidelberg.^[3] In 1905 he was the president of the [Deutsche Mathematiker-Vereinigung](#). His doctoral students include [Paul Riebesell](#).



Paul Stäckel.

Works [[edit](#)]

- *Über die Bewegung eines Punktes auf einer Fläche*^[?], 1885, Dissertation
- *Die Integration der Hamilton-Jacobischen Differentialgleichung mittelst Separation der Variablen*, 1891, Habilitation
- *Elementare Dynamik der mathematischen Wissenschaft* in: Encyclopädie der mathematischen Wissenschaft IV,1 (1908)
- several papers in the *Mathematischen Annalen* (from 1890 to 1909)
- numerous papers in "Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse" from 1896 to 1917.
- with [Friedrich Engel](#): *Theorie der Parallellinien von Euklid bis auf Gauss*^[?], 1895

Text

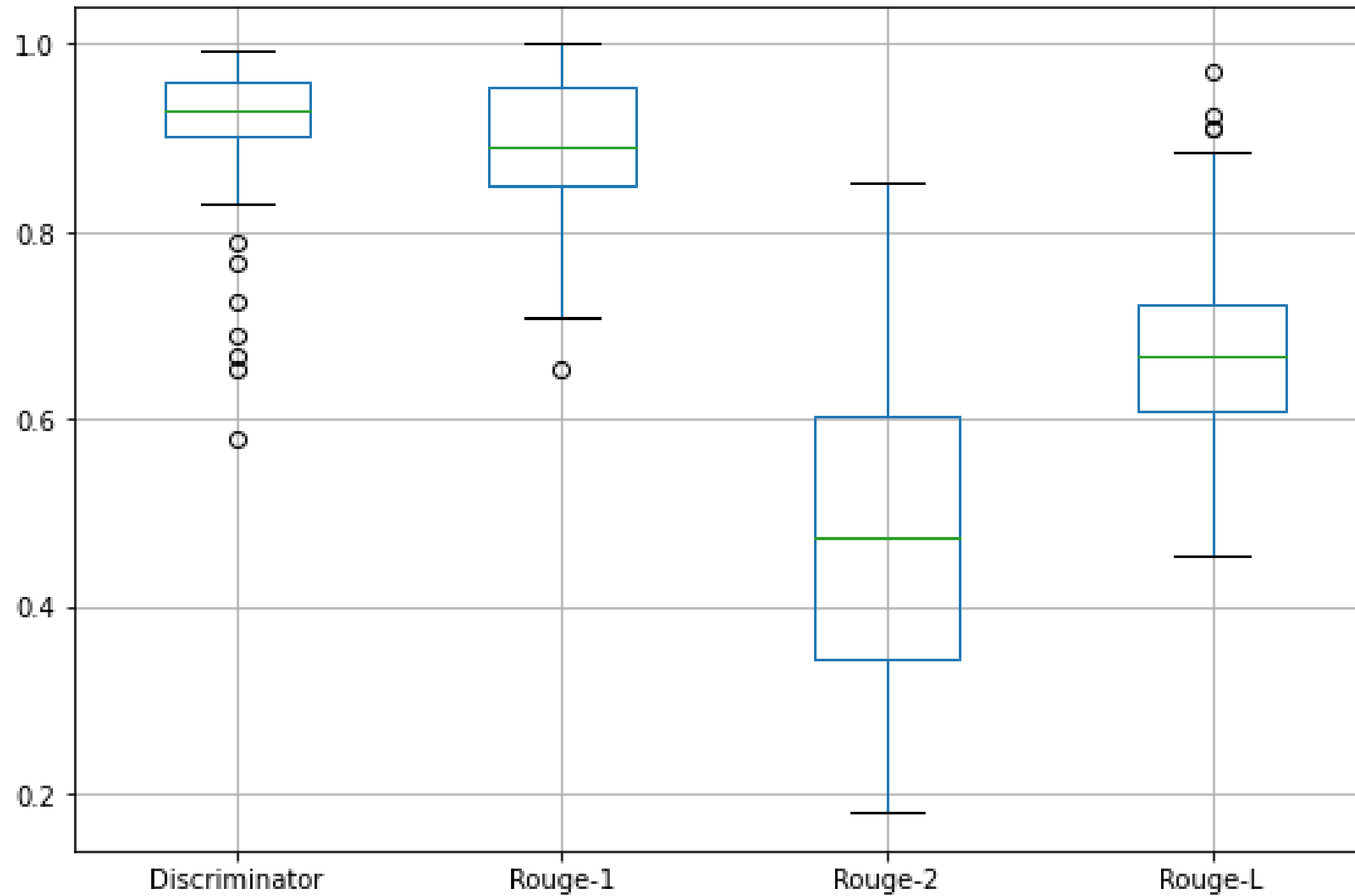
Results

1

Discriminator

Discriminator results on CNN/DailyMail

It works!



We also created our own (small) dataset

Goal

Evaluate different aspects, with special emphasis in incorrect summaries.

Our own dataset – Categories

- Good summaries:
 - Ground truth.

Our own dataset – Categories

- Good summaries:
 - Ground truth.
 - Synonyms: The ground-truth summary replacing words with synonyms.

Our own dataset – Categories

- Good summaries:
 - Ground truth.
 - Synonyms: The ground-truth summary replacing words with synonyms.
- Bad summaries:
 - Unrelated: A summary that doesn't correspond with the original text.

Our own dataset – Categories

- Good summaries:
 - Ground truth.
 - Synonyms: The ground-truth summary replacing words with synonyms.
- Bad summaries:
 - Unrelated: A summary that doesn't correspond with the original text.
 - Nonsensical: A summary that is related to original text but makes no sense semantically.

Our own dataset – Categories

- Good summaries:
 - Ground truth.
 - Synonyms: The ground-truth summary replacing words with synonyms.
- Bad summaries:
 - Unrelated: A summary that doesn't correspond with the original text.
 - Nonsensical: A summary that is related to original text but makes no sense semantically.
 - Keywords: A summary that includes a lot of words from the original text, but misses the point of it.

Our own dataset – Categories

- Good summaries:
 - Ground truth.
 - Synonyms: The ground-truth summary replacing words with synonyms.
- Bad summaries:
 - Unrelated: A summary that doesn't correspond with the original text.
 - Nonsensical: A summary that is related to original text but makes no sense semantically.
 - Keywords: A summary that includes a lot of words from the original text, but misses the point of it.
 - Bad grammar: A summary that could be appropriate but contains several grammar mistakes.

Our own dataset – Categories

- Good summaries:
 - Ground truth.
 - Synonyms: The ground-truth summary replacing words with synonyms.
- Bad summaries:
 - Unrelated: A summary that doesn't correspond with the original text.
 - Nonsensical: A summary that is related to original text but makes no sense semantically.
 - Keywords: A summary that includes a lot of words from the original text, but misses the point of it.
 - Bad grammar: A summary that could be appropriate but contains several grammar mistakes.
 - Negations: A correct summary but negated, so that it conveys the opposite meaning.

Evaluation Scores		Discriminator	Rouge-1	Rouge-2	Rouge-L
Own dataset					
Good Summaries	Ground Truth	0.96	0.88	0.64	0.78
	Synonyms	0.96	0.70	0.37	0.56
Bad Summaries	Unrelated	0.21	0.26	0.01	0.18
	Nonsensical	0.90	0.84	0.21	0.51
	Keywords	0.58	0.68	0.24	0.58
	Bad grammar	0.95	0.82	0.48	0.69
	Negations	0.94	0.68	0.41	0.59

Discriminator caveats

Currently **difficult** to **generalize** to unseen dataset.

- ▶ More data and training should help.

Discriminator caveats

Currently **difficult** to **generalize** to unseen dataset.

► More data and training should help.

Limitations in **resources**: GPU RAM \Rightarrow Small batch size \Rightarrow No training of batchnorm.

► Having higher compute would not only be faster but also better.

Discriminator caveats

Currently **difficult** to **generalize** to unseen dataset.

- ▶ More data and training should help.

Limitations in **resources**: GPU RAM \Rightarrow Small batch size \Rightarrow No training of batchnorm

- ▶ Having higher compute would not only be faster but also better.

Limitations in **work capacity**.

- ▶ Could create a dataset targeting things like negations or synonyms, but would take too long.

2

GAN + Generator

GAN + Generator results

Pipeline running but some problems:

- For each iteration, a good GPU (16 GB) needs a day.
 - GAN loop slightly unstable.
- Until now generator not improving on pretrained base.

.

Abstractive Summarization

Wikipedia (+ CNN/DailyMail)

Thank you for your attention.

