Paper Review

# Learning to Perturb Word Embeddings for Out-of-distribution QA

**Lee et al., ACL, 2021**


**Myeongsup Kim**

Integrated M.S./Ph.D. Student
Data Science & Business Analytics Lab.
School of Industrial Management Engineering
Korea University

Myeongsup_kim@korea.ac.kr

# <What Are Not Covered in This Presentation>

- **Details of Regularization**

  Goodfellow et al., 2016, Deep Learning, MIT Press, Chapter 7

- **Details of Transformer**

  Vaswani et al., 2017, Attention is All You Need, NIPS

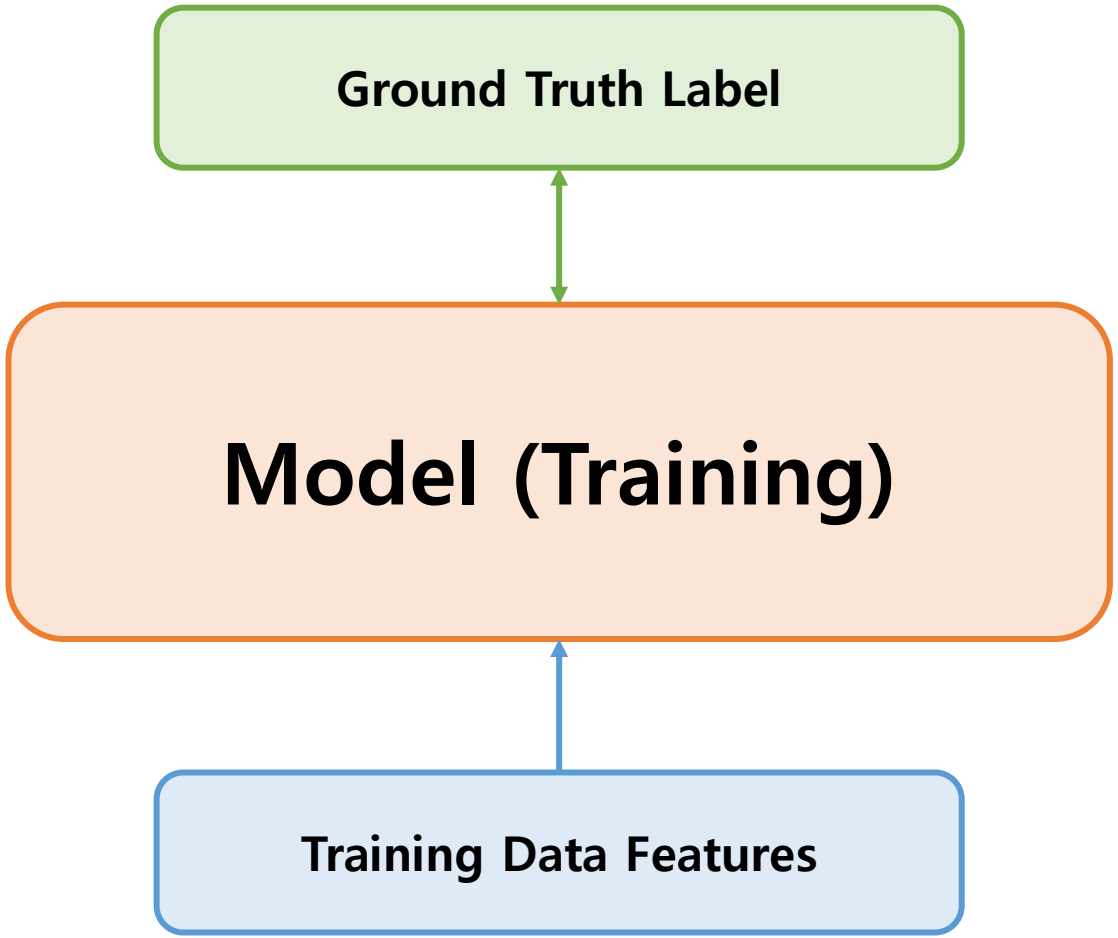
- **Details of BERT**

  Devlin et al., 2019, BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding, NAACL

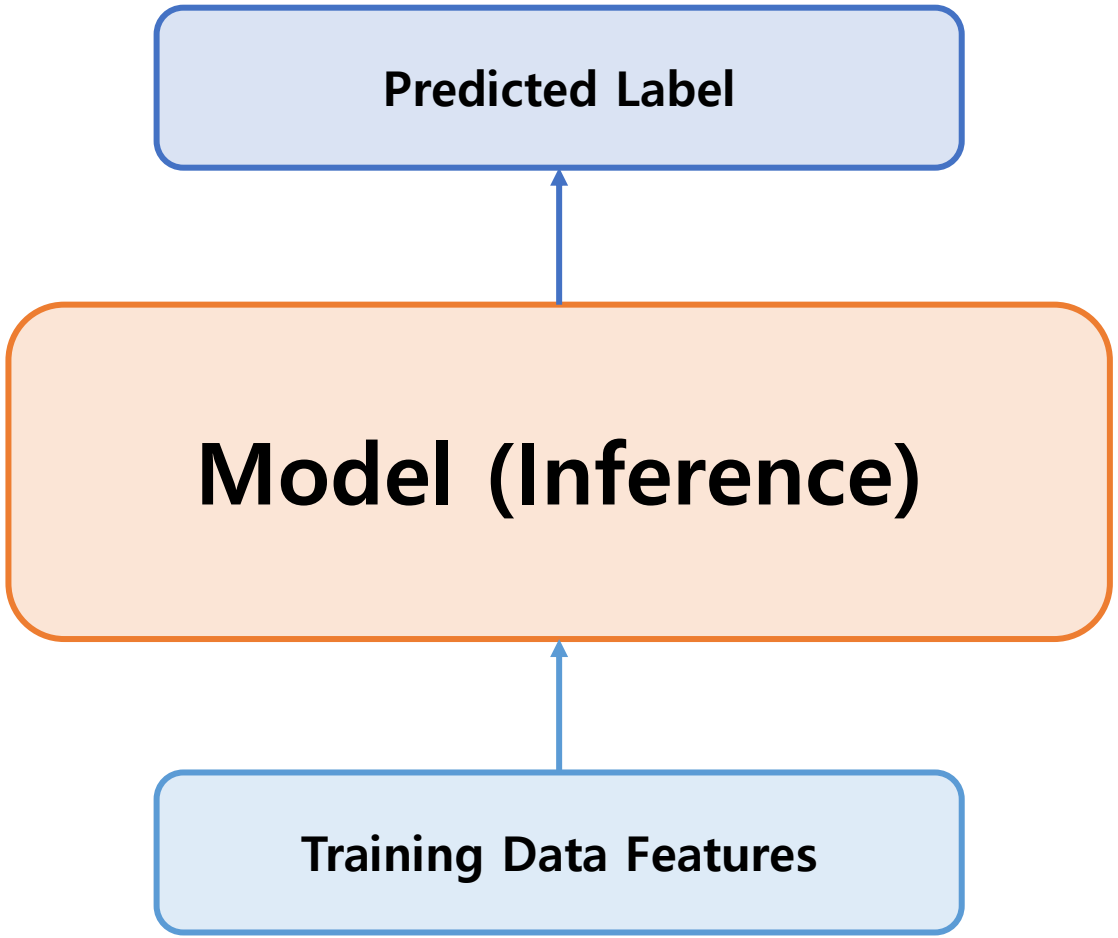
# Introduction

- Text Data Augmentation
- Question Answering

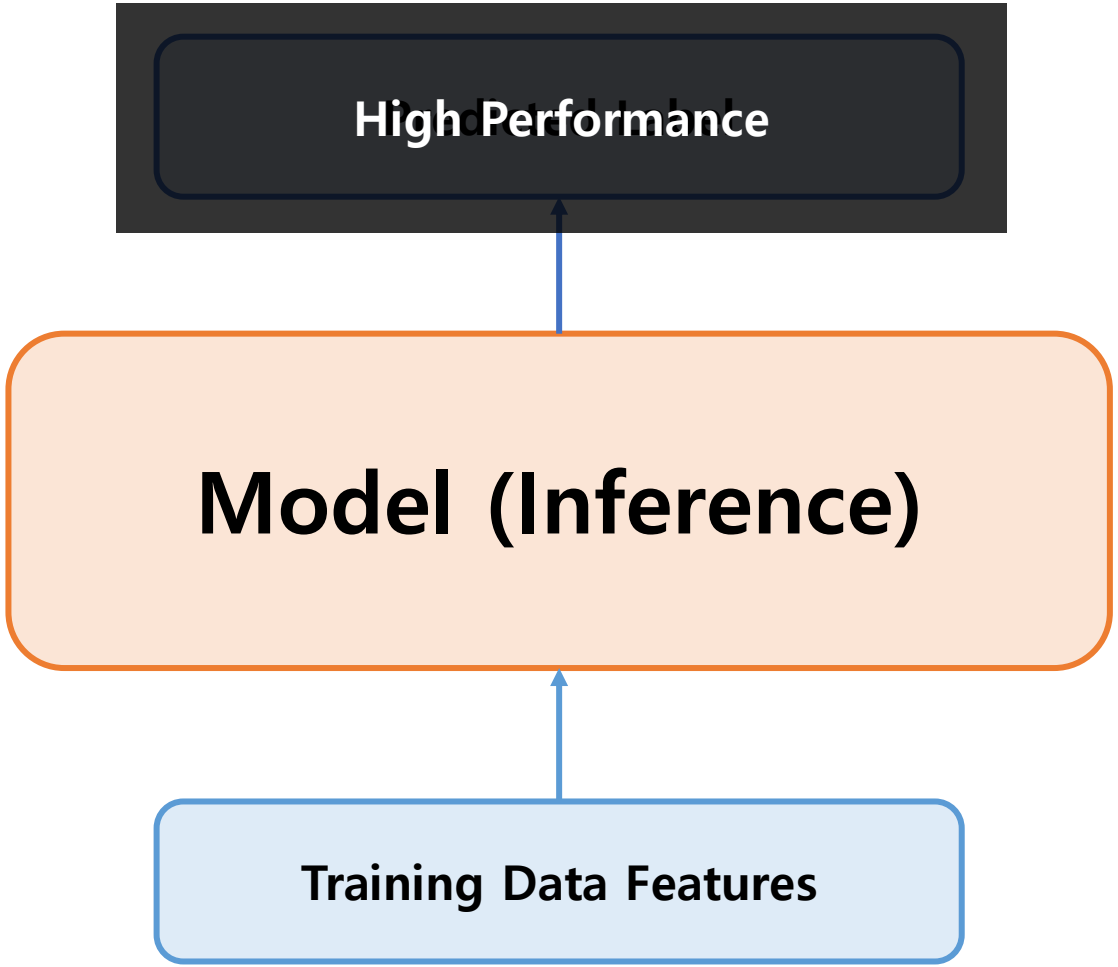
## <Supervised Learning>



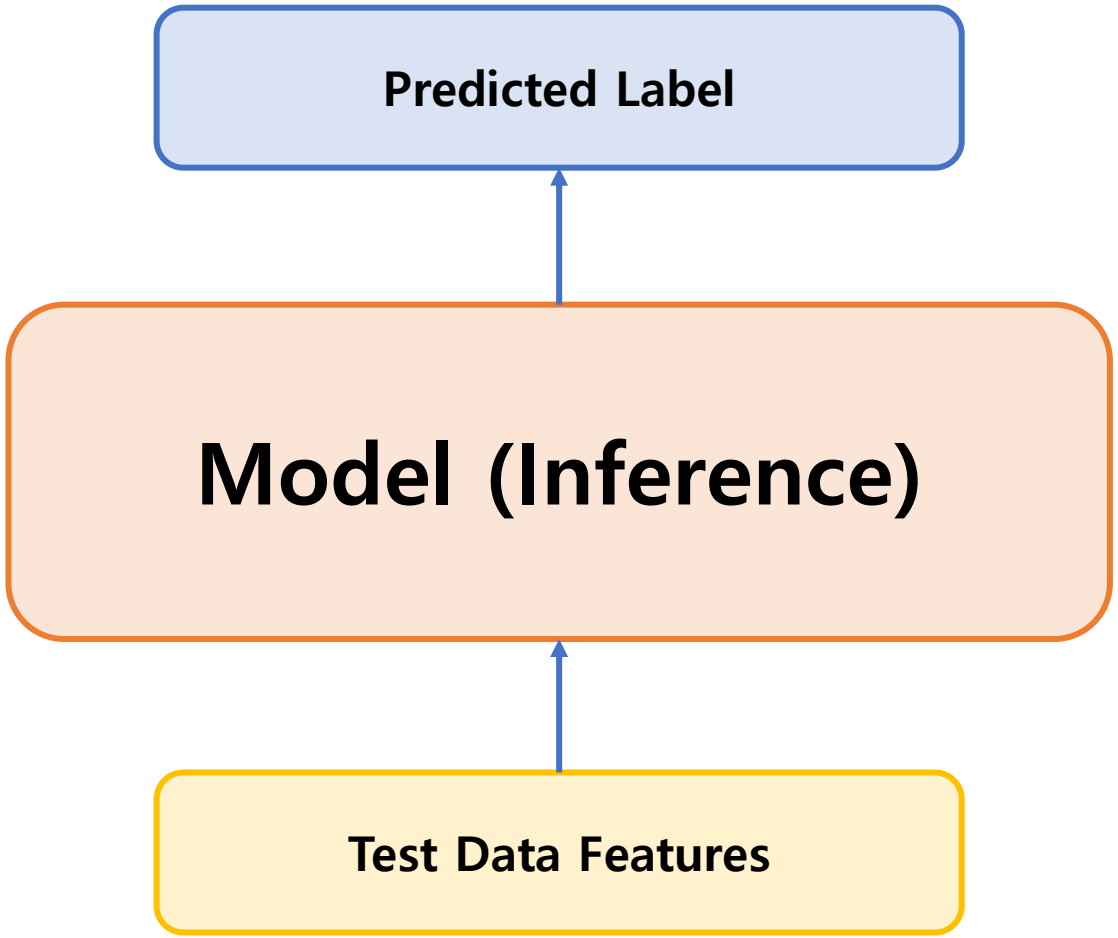
## <Overfitting>



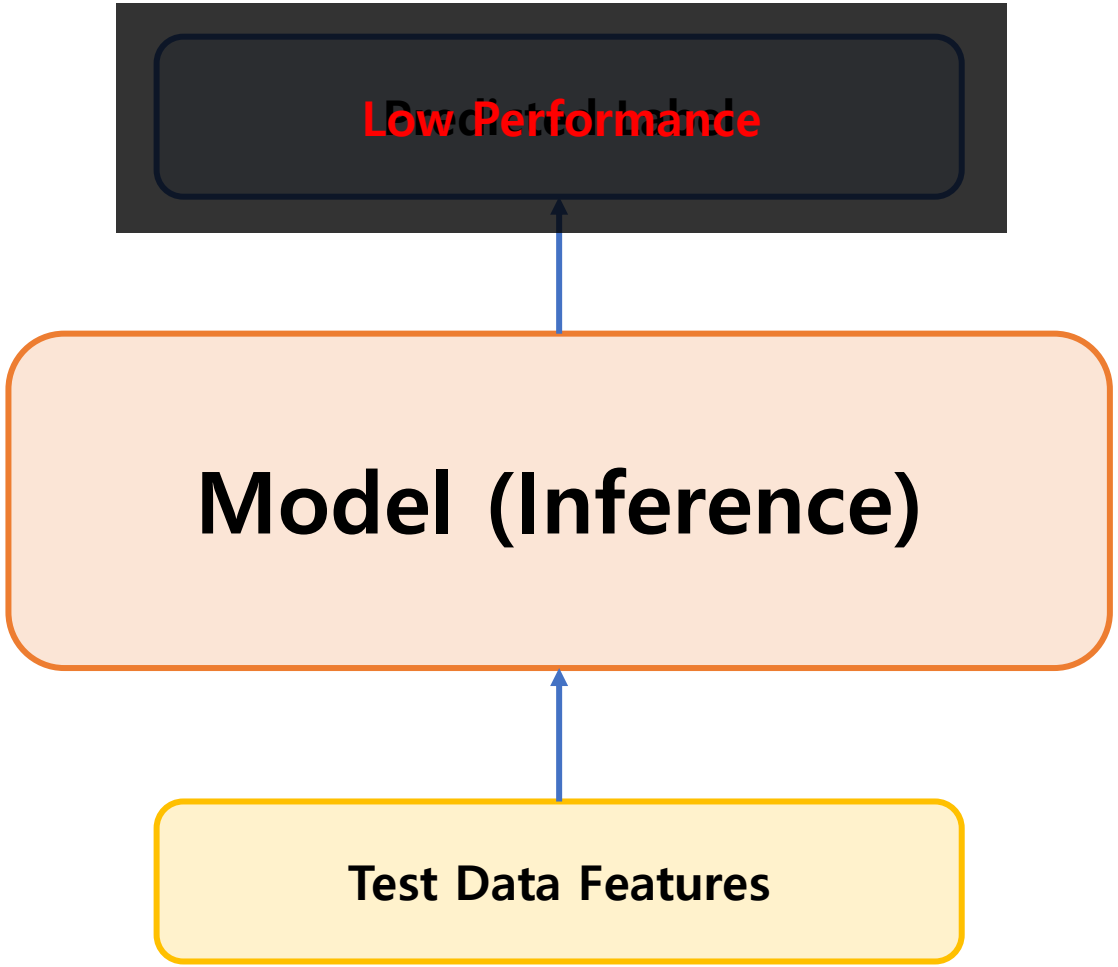
## <Overfitting>



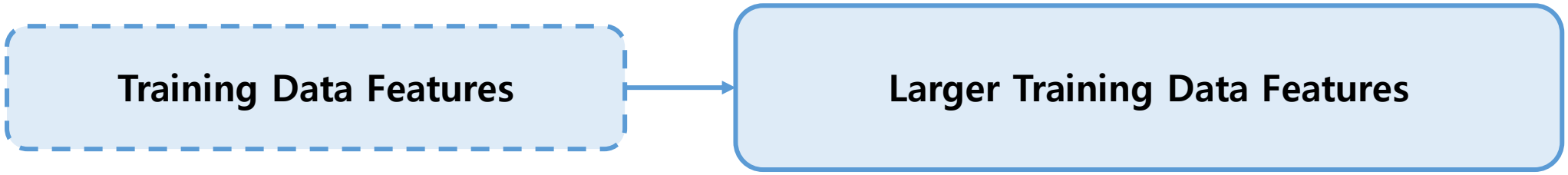
## <Overfitting>



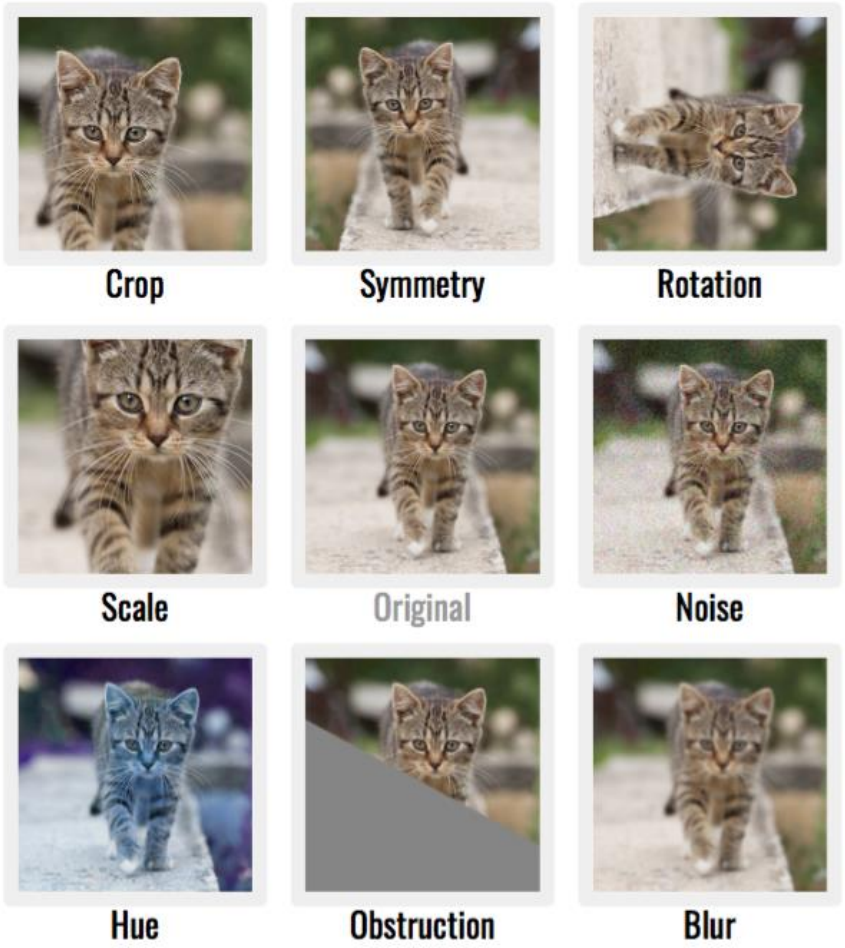
## <Overfitting>



# <Data Augmentation>



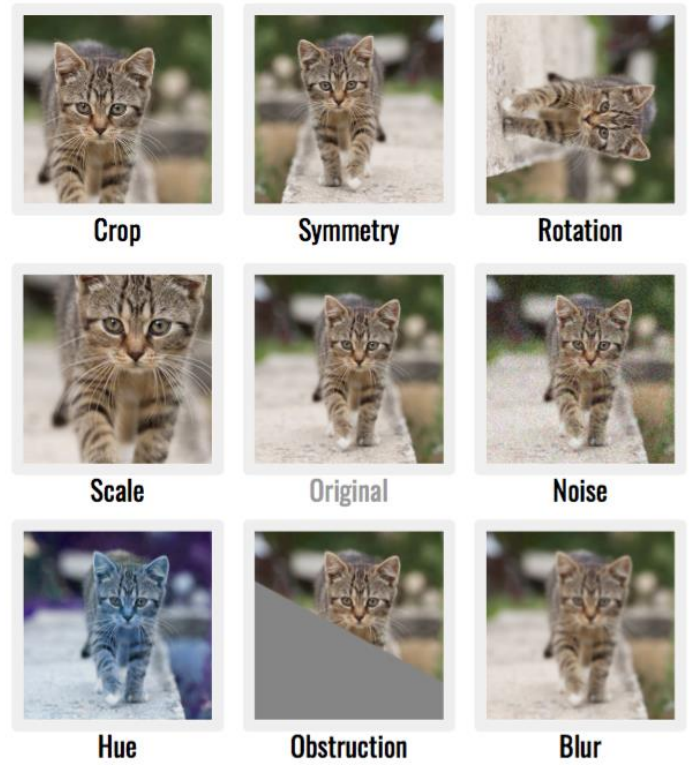
# <Data Augmentation>



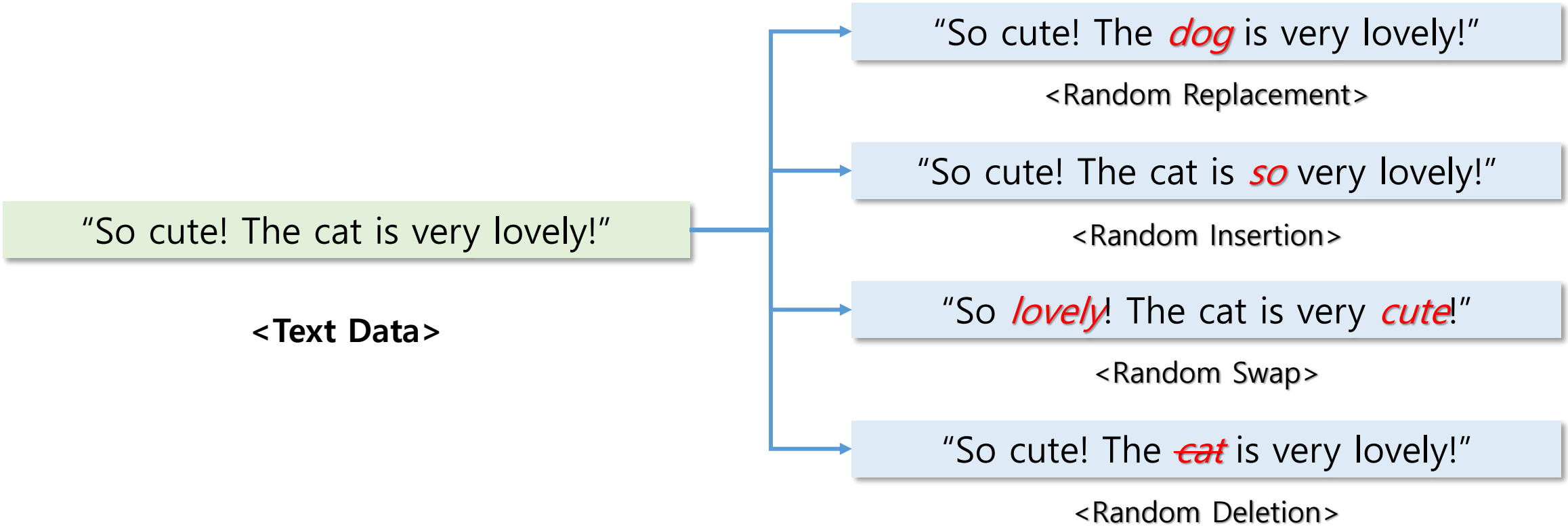
# <Text Data Augmentation>

"So cute! The cat is very lovely!"

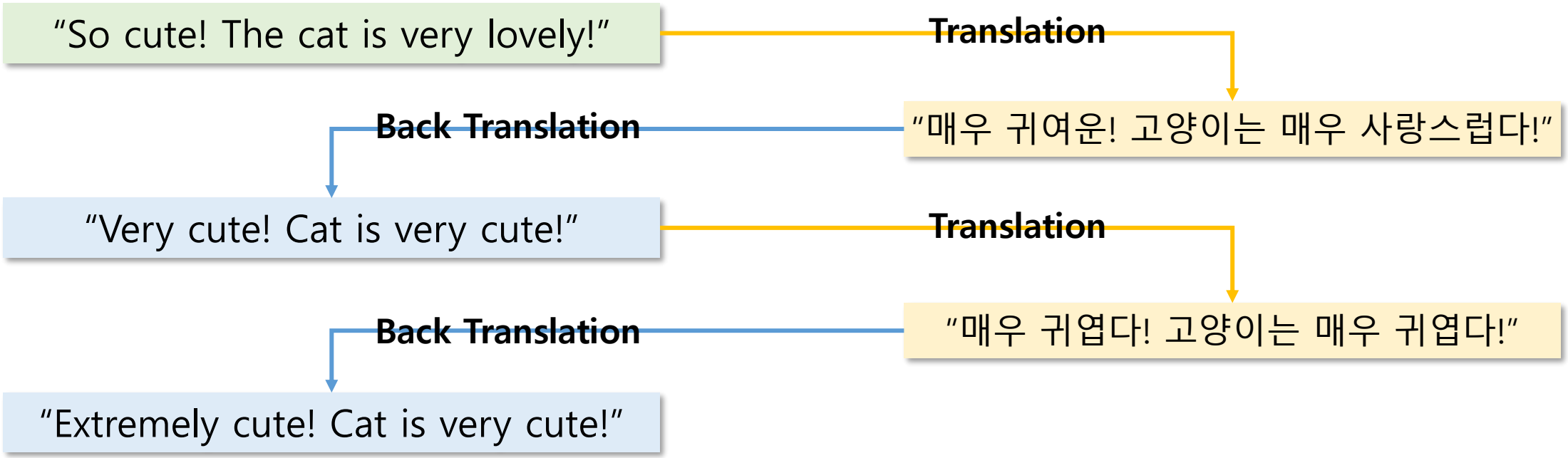
**<Text Data>**



# <Easy Data Augmentation>



# <Back Translation>



# <Question Answering>

## Passage

The Normans (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11 th centuries gave their name to Normandy, a region in France. They were descended from Norse ("Norman" come s from "Norseman") raiders and pirates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Francia. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants would gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Normans emerged initially in the fir st half of the 10th century, and it continued to evolve over the succeeding centuries.

## Question 1

In what country is Normandy located?

## Question 2

When were the Normans in Normandy?

# <Question Answering>

## Passage

The Normans (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11 th centuries gave their name to Normandy, a region in France. They were descended from Norse ("Norman" come s from "Norseman") raiders and pirates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Francia. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants would gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Normans emerged initially in the fir st half of the 10th century, and it continued to evolve over the succeeding centuries.

## Question 1

In what country is Normandy located?

## Question 2

When were the Normans in Normandy?

## Answer 1

France

## Answer 2

$10^{th}$ and $11^{th}$ centuries

# <Text Augmentation for QA>

## Passage

The ~~**Normans**~~ (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11th centuries gave their name to Normandy, a region in **Germany**. They were descended from Norse ("Norman" comes from "Norseman") raiders and pirates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Francia. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants would gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Normans emerged initially in the first half of the 10th century, and it continued to evolve over the succeeding centuries.

# <Text Augmentation for QA>

## Passage

The ~~Normans~~ (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11th centuries gave their name to Normandy, a region in Germany. They were descended from Norse ("Norman" comes from "Norseman") raiders and pirates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Francia. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants would gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Normans emerged initially in the first half of the 10th century, and it continued to evolve over the succeeding centuries.

## Question 1

In what country is Normandy located?

## Question 2

When were the Normans in Normandy?

# <Text Augmentation for QA>

## Passage

The ~~Normans~~ (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11th centuries gave their name to Normandy, a region in **Germany**. They were descended from Norse ("Norman" comes from "Norseman") raiders and pirates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Francia. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants would gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Normans emerged initially in the first half of the 10th century, and it continued to evolve over the succeeding centuries.

### Question 1

In what country is Normandy located?

### Answer 1

**Germany**

### Question 2

When were the Normans in Normandy?

### Answer 2

**Impossible to answer**

# <Text Augmentation for QA>

**Using Token Level Text Augmentation for Question Answering Task May Change the Answer for a Question.**

**How Can We Augment Text Data for Question Answering Task?**

Passage

The Normans (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11th centuries gave their name to Normandy, a region in France. They were descended from Norse ("Norseme n") raiders and pirates from Denmark, Iceland and Norway who, under their leader Rollo, agr eed to swear fealty to King Charles III of West Francia. Through generations of assimilation and mixing with the na tive Frankish and Roman-Gaulish populations, their descendants would gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Normans emerged initially in the fir st half of the 10th century, and it continued to evolve over the succeeding centuries.

Question 1

In what country is Normandy located?

Answer 1

Germany

Question 2

When were the Normans in Normandy?

Answer 2

Impossible to answer

# Learning to Perturb Word Embeddings for Out-of-distribution QA

*Lee et al., ACL, 2021*

# Learning to Perturb Word Embeddings for Out-of-distribution QA

*Lee et al., ACL, 2021*

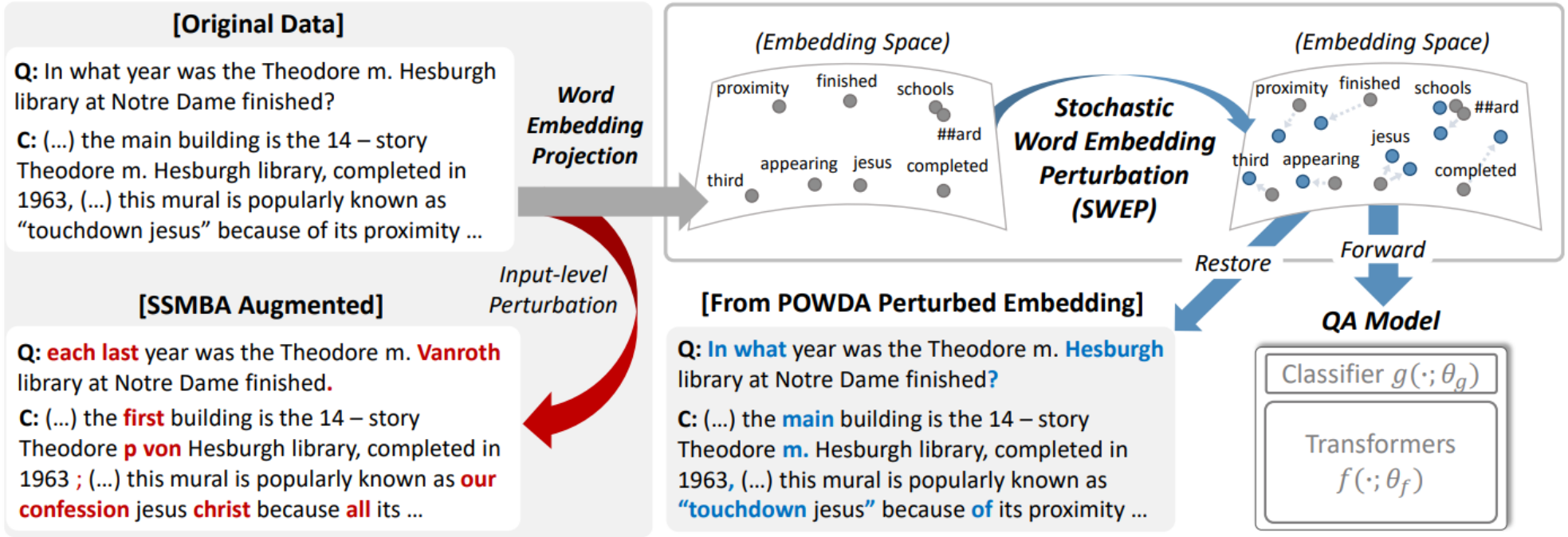
# SWEP

- Question Answering Model
- Learning to Perturb Word Embedding
- Learning Objective

# Method

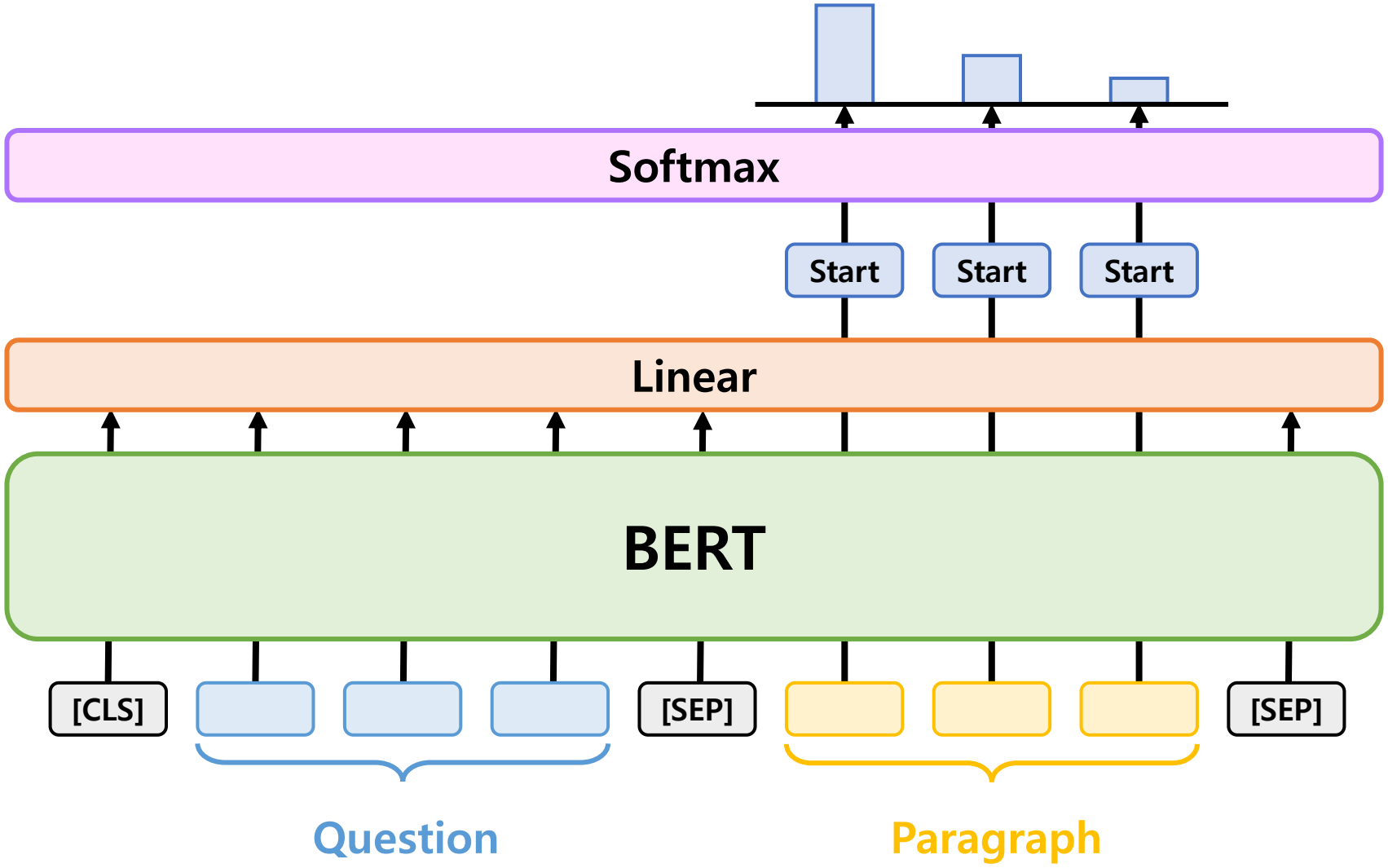
-Concept

## <Overall Concept>



# Method

## -Question Answering Model



# <Question Answering Model>

## Passage

The Normans (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11 th centuries gave their name to Normandy, a region in France. They were descended from Norse ("Norman" come s from "Norseman") raiders and pirates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Francia. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants would gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Normans emerged initially in the fir st half of the 10th century, and it continued to evolve over the succeeding centuries.

## Question

When were the Normans in Normandy?

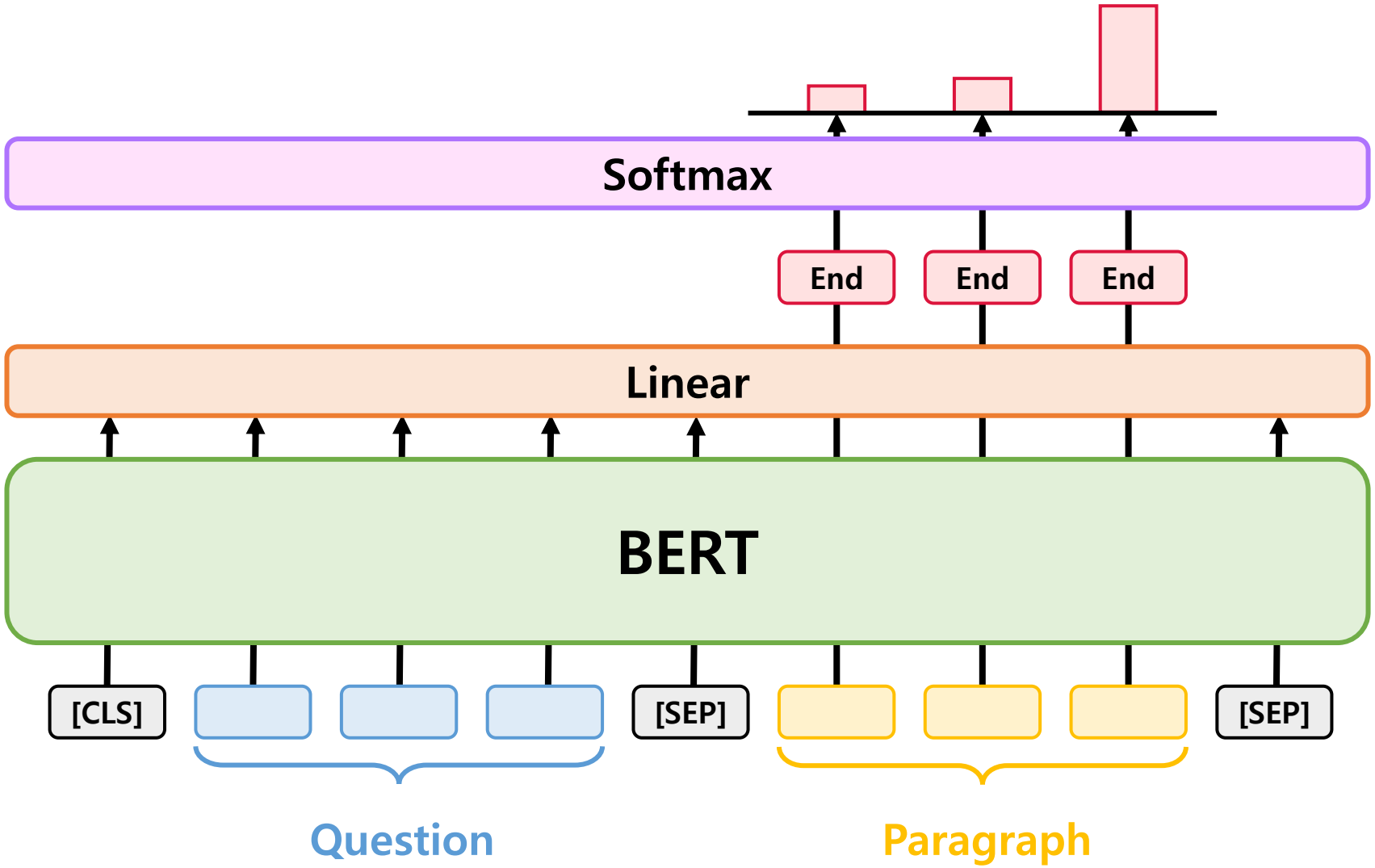
## Answer

$10^{th}$ and $11^{th}$ centuries

# Method

## -Question Answering Model

<Question Answering Model>



# <Question Answering Model>

## Passage

The Normans (Norman: Nourmands; French: Normands; Latin: Normanni) were the people who in the 10th and 11 th centuries gave their name to Normandy, a region in France. They were descended from Norse ("Norman" come s from "Norseman") raiders and pirates from Denmark, Iceland and Norway who, under their leader Rollo, agreed to swear fealty to King Charles III of West Francia. Through generations of assimilation and mixing with the native Frankish and Roman-Gaulish populations, their descendants would gradually merge with the Carolingian-based cultures of West Francia. The distinct cultural and ethnic identity of the Normans emerged initially in the fir st half of the 10th century, and it continued to evolve over the succeeding centuries.

## Question

When were the Normans in Normandy?

## Answer

10th and 11th centuries

# <Objective Function>

$$\mathcal{L}_{MLE}(\theta) := \sum_{i=1}^{N} \log p_\theta \left(\mathbf{y}^{(i)} \mid \mathbf{x}^{(i)}, \mathbf{c}^{(i)}\right)$$

$\mathbf{y} = (y_{start}, y_{end})$ : $Answer\ Span$

$\mathbf{x} = (x_1, \ldots, x_M)$ : $Question$

$\mathbf{c} = (c_1, \ldots, c_L)$ : $Context$

$f(\cdot;\ \theta_f)$ : $Encoder, Parameters$

$g(\cdot;\ \theta_g)$ : $Decoder(or\ Classifier), Paramters$

$p_\theta(\mathbf{y}|\mathbf{x}, \mathbf{c}) = g(f(\mathbf{y}|\mathbf{x}, \mathbf{c}; \theta_f); \theta_g)$ : $Model$

# <Objective Function>

$$\mathcal{L}_{MLE}(\theta) := \sum_{i=1}^{N} \log p_\theta \left(\mathbf{y}^{(i)} \mid \mathbf{x}^{(i)}, \mathbf{c}^{(i)}\right)$$

$\mathbf{y} = (y_{start}, y_{end})$ : *Answer Span*

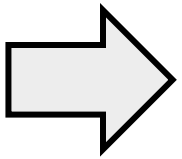
~~$\mathbf{x} = (x_1, \dots, x_M)$ : *Question*~~

~~$\mathbf{c} = (c_1, \dots, c_L)$ : *Context*~~

$f(\cdot;\ \theta_f)$ : *Encoder, Parameters*

$g(\cdot;\ \theta_g)$ : *Decoder(or Classifier), Paramters*

$p_\theta(\mathbf{y}|\mathbf{x}, \mathbf{c}) = g(f(\mathbf{y}|\mathbf{x}, \mathbf{c}; \theta_f); \theta_g)$ : *Model*



$$\mathcal{L}_{MLE}(\theta) := \sum_{i=1}^{N} \log p_\theta \left(\mathbf{y}^{(i)} \mid \mathbf{x}^{(i)}\right)$$

$\mathbf{y} = (y_{start}, y_{end})$ : *Answer Span*

$\mathbf{x} = (x_0, x_1, \dots, x_M, c_0, \dots, c_{M+1})$ : *Input*

$T = L + M + 3$ : *Input Lenght*

$f(\cdot;\ \theta_f)$ : *Encoder, Parameters*

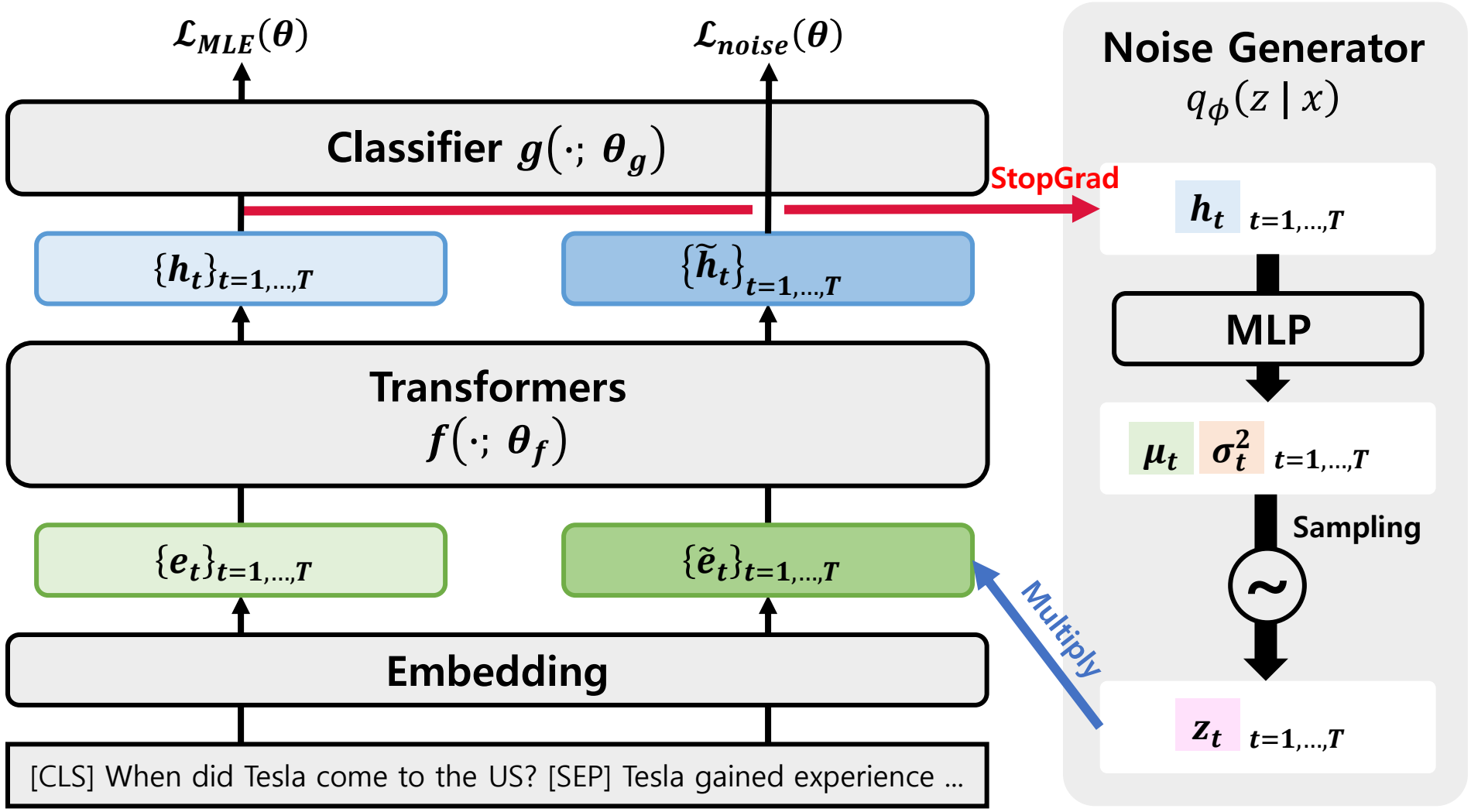
$g(\cdot;\ \theta_g)$ : *Decoder(or Classifier), Paramters*

$p_\theta(\mathbf{y}|\mathbf{x}, \mathbf{c}) = g(f(\mathbf{y}|\mathbf{x}, \mathbf{c}; \theta_f); \theta_g)$ : *Model*

# Method

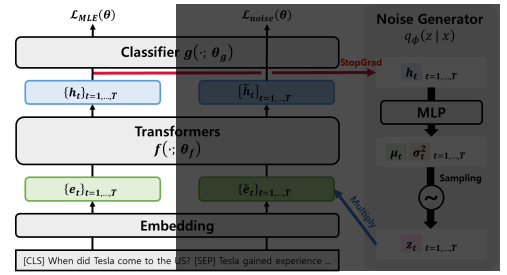
-Learning to Perturb Word Embedding

## <Stochastic Word Embedding Perturbation>

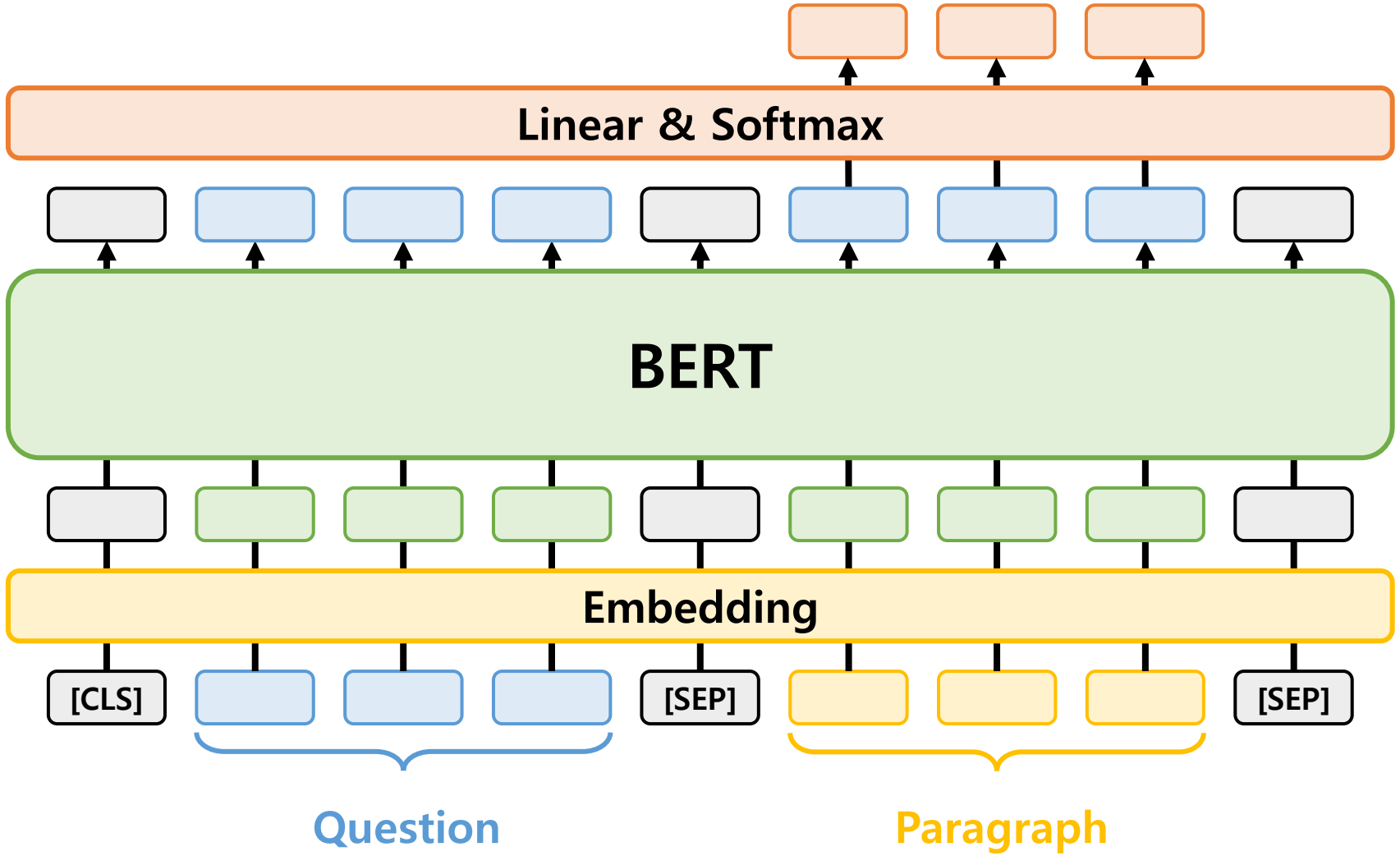


# Method

-Learning to Perturb Word Embedding

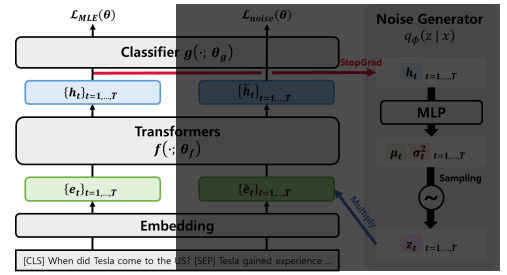



## <Extract Representation>



# Method

## -Learning to Perturb Word Embedding




## <Extract Representation>

Stop Gradient

BERT

Embedding

[CLS] [SEP] [SEP]

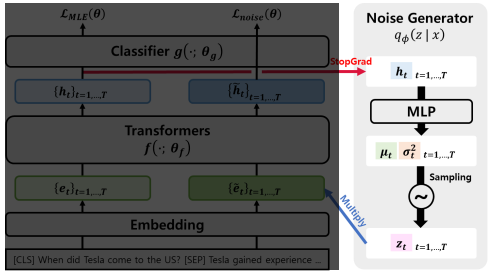
Noise Generator

Question

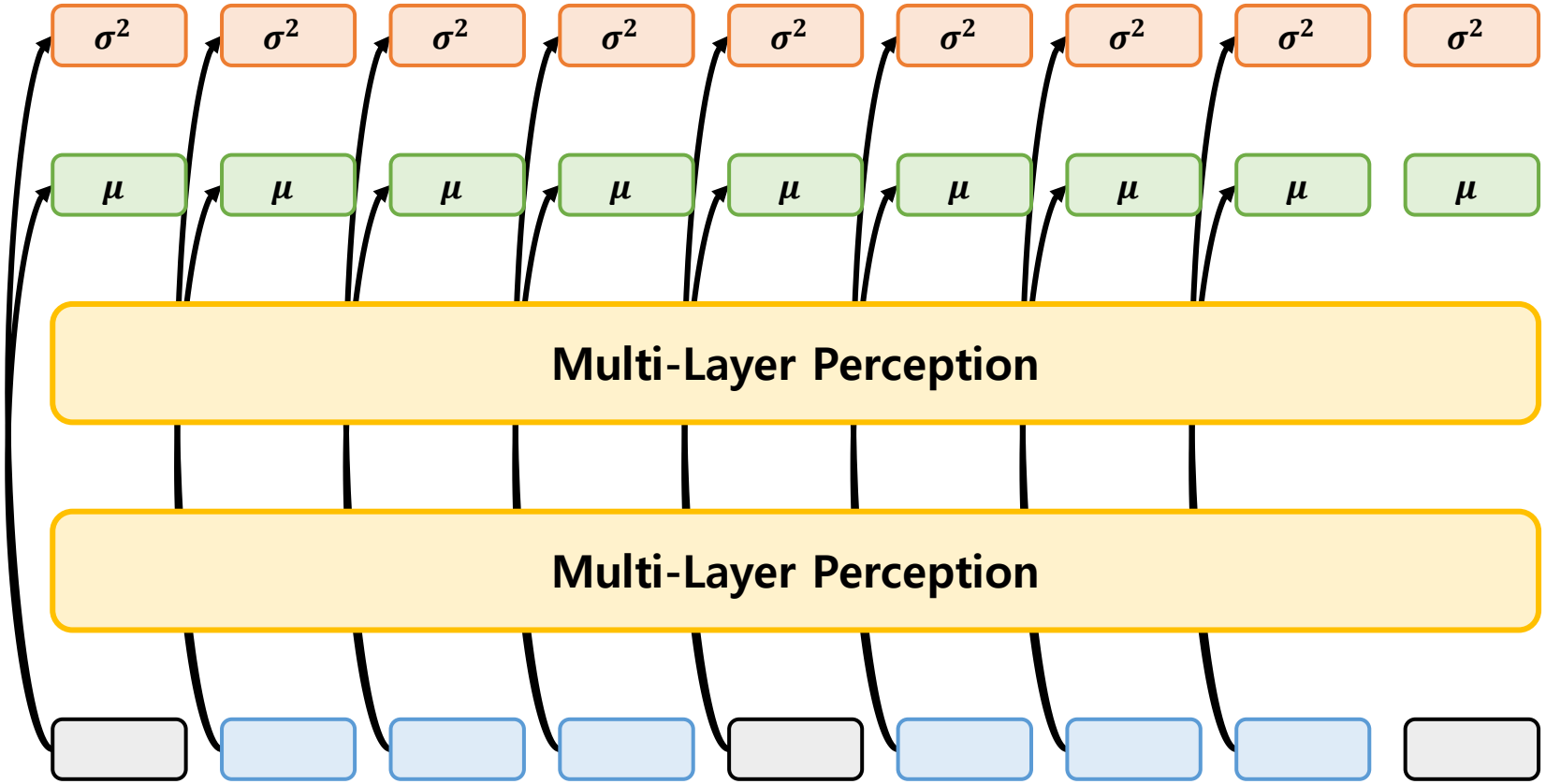
Paragraph

# Method

## -Learning to Perturb Word Embedding

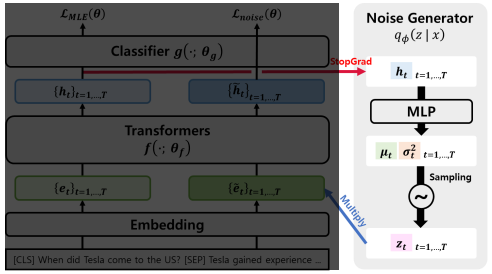



## <Noise Generator>

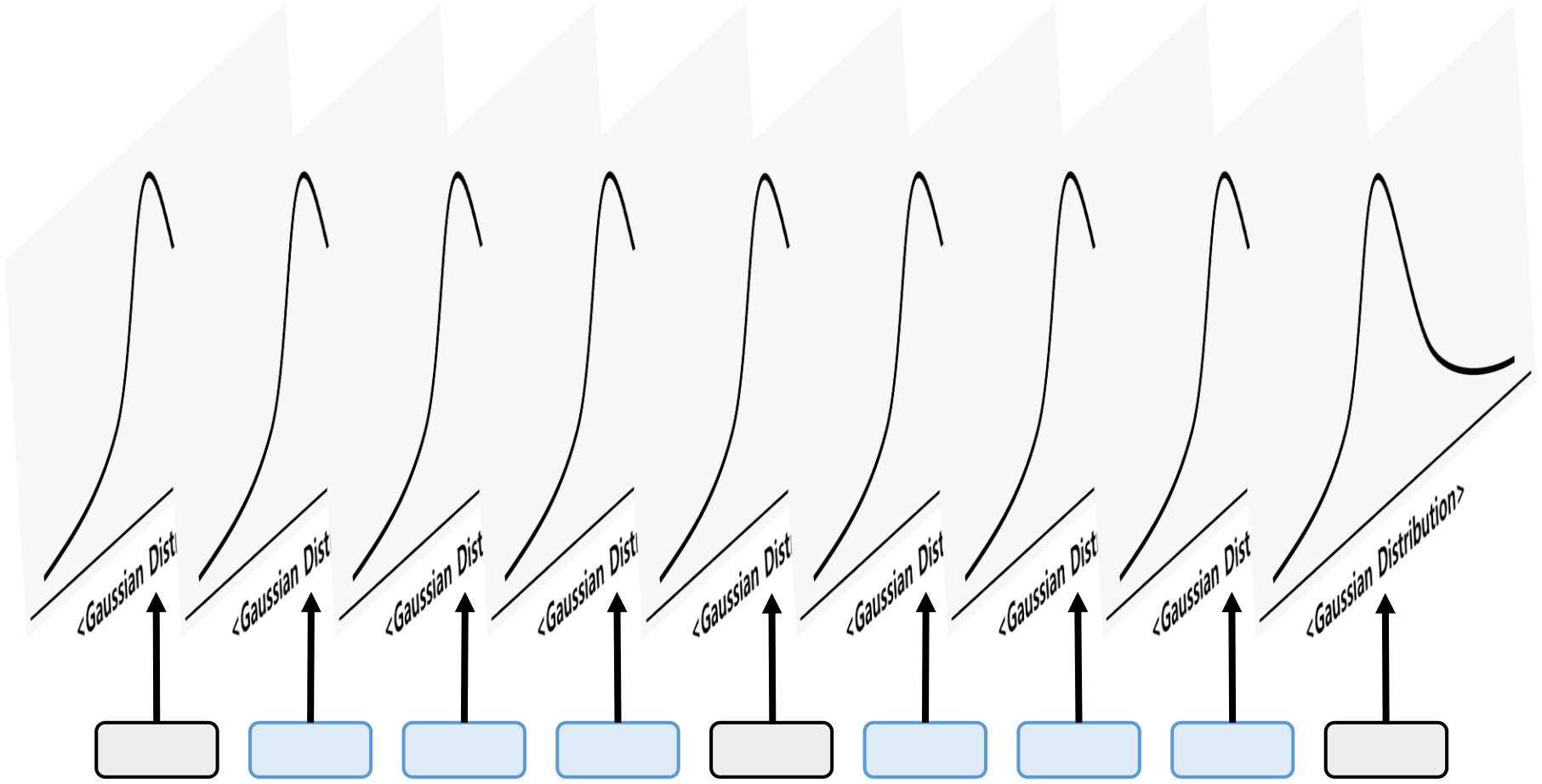


# Method

## -Learning to Perturb Word Embedding

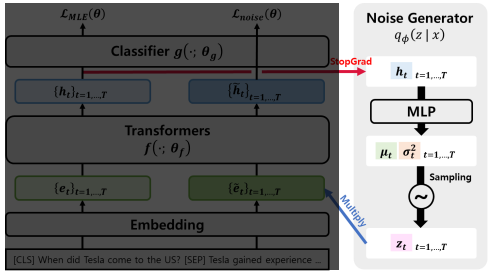



# <Noise Generator>



# Method

## -Learning to Perturb Word Embedding




## <Noise Generator>

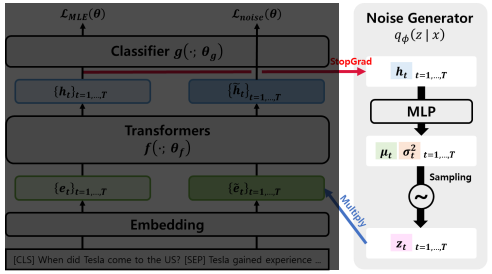
Sampling

z z z z z z z z z

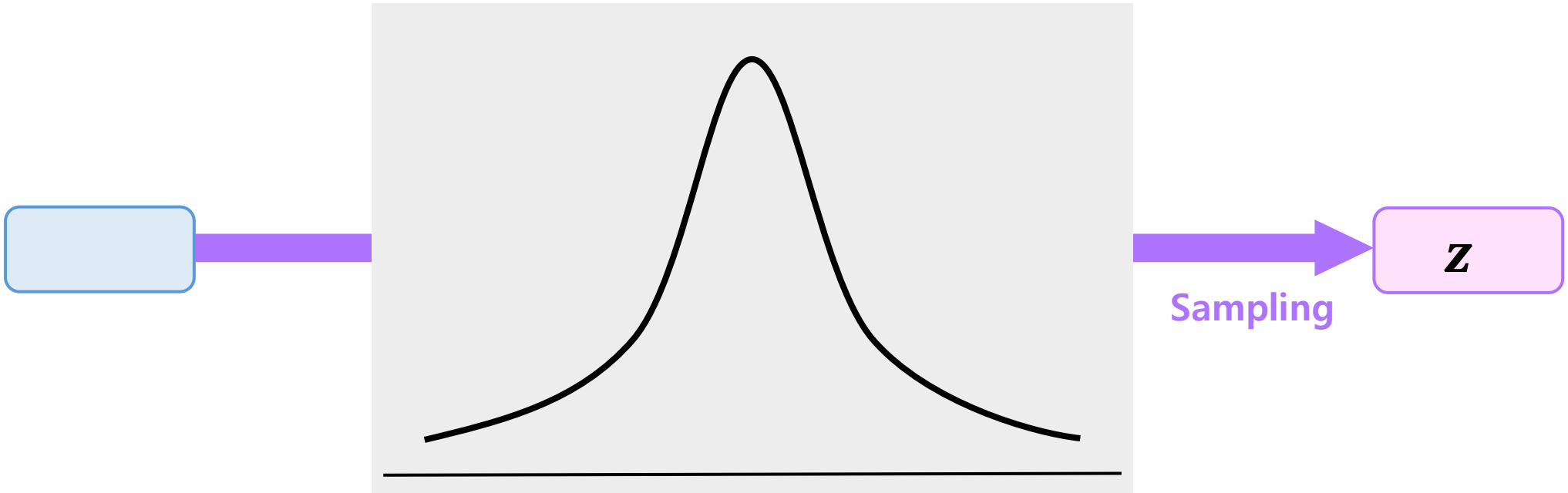
<Gaussian Distribution>

# Method

## -Learning to Perturb Word Embedding



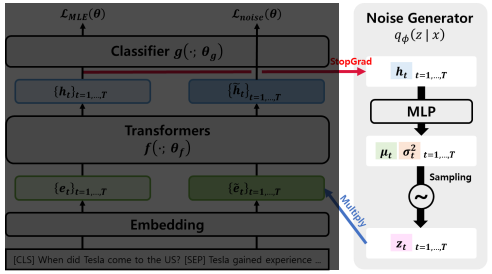

# <Reparameterization Trick>



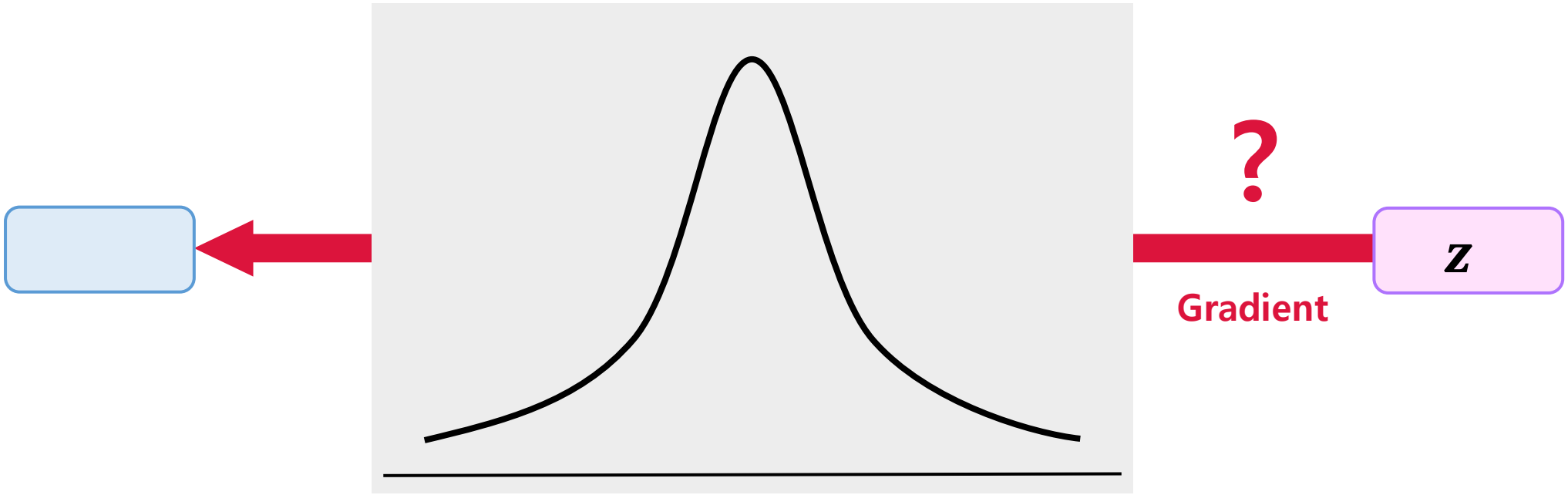

<Gaussian Distribution>

# Method

## -Learning to Perturb Word Embedding



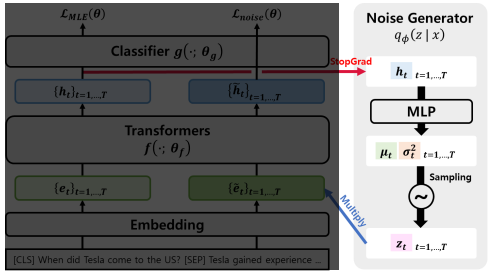

# <Reparameterization Trick>




<Gaussian Distribution>

# Method

## -Learning to Perturb Word Embedding




## <Reparameterization Trick>

$$\mathbf{e}_t = WordEmbedding(x_t)$$

$$(\mathbf{h}_1, \dots, \mathbf{h}_T) = f(\mathbf{e}_1, \dots, \mathbf{e}_T;\ \theta_f)$$

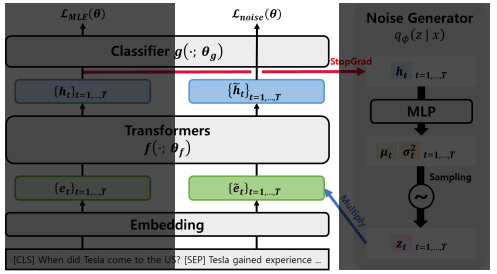
$$\boldsymbol{\mu}_t, \boldsymbol{\sigma}_t^2 = \mathbf{MLP}(\boldsymbol{h}_t)$$

$$\boldsymbol{z}_t = \boldsymbol{\mu}_t + \boldsymbol{\sigma}_t^2 \ \odot\ \epsilon, where\ \epsilon \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_d)$$

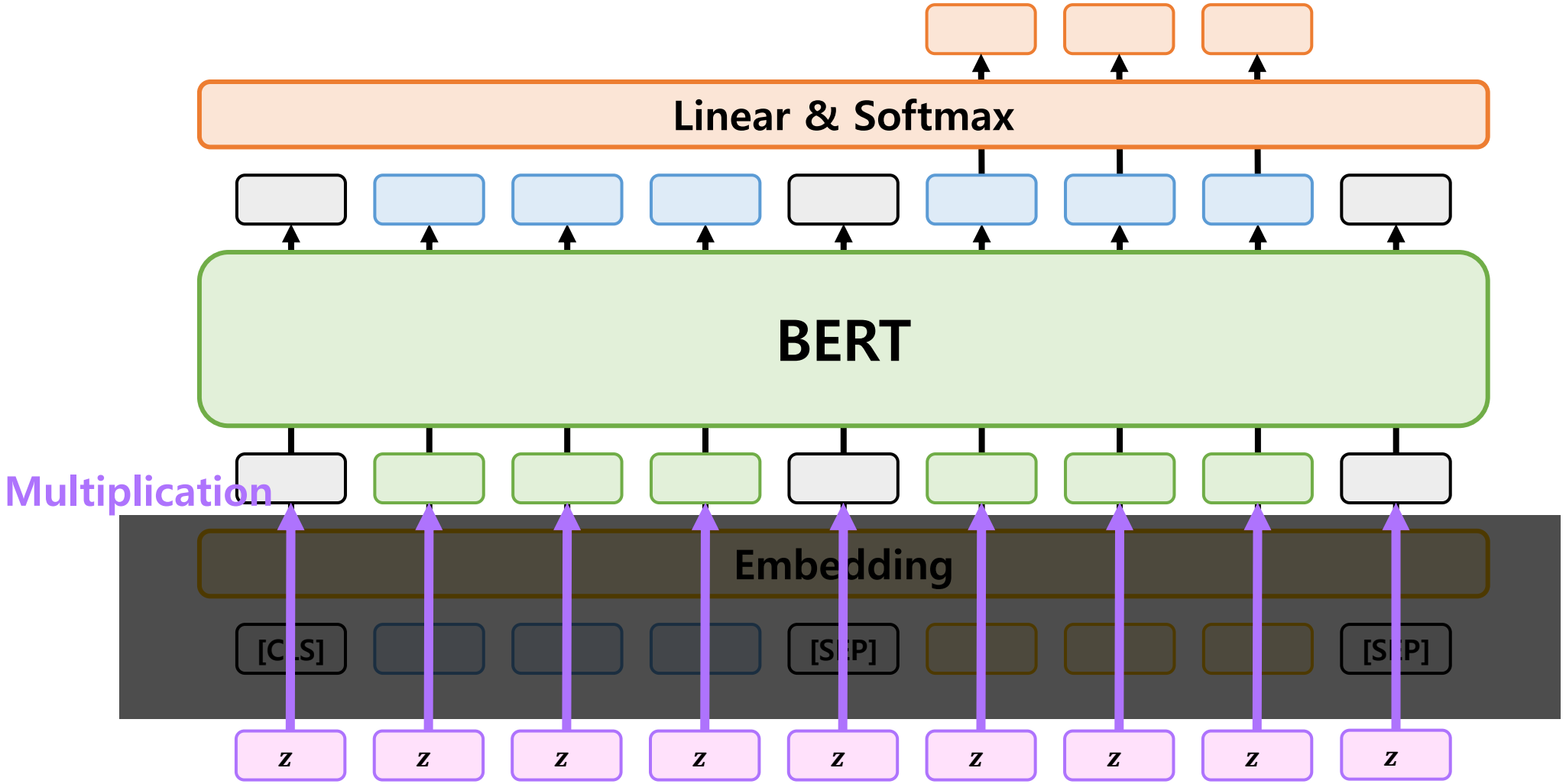
$$\tilde{\mathbf{e}}_t = \ \mathbf{e}_t + \boldsymbol{z}_t$$

# Method

## -Learning to Perturb Word Embedding




# <Learning to Perturb Embedding>



# <Learning Objective>

$$\mathcal{L}(\phi, \theta) = \lambda \mathcal{L}_{MLE}(\theta) + (1 - \lambda)\mathcal{L}_{noise}(\phi, \theta)$$

$$\mathcal{L}_{MLE}(\theta) \coloneqq \sum_{i=1}^{N} \log p_\theta \left(\mathbf{y}^{(i)} \mid \mathbf{x}^{(i)}\right)$$

$$\mathcal{L}_{noise}(\phi, \theta) \coloneqq \sum_{i=1}^{N} \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x}^{(i)})}\left[\log_{p_\theta}\left(\mathbf{y}^{(i)}|\mathbf{x}^{(i)}, \mathbf{z}\right)\right]$$
$$-\beta \sum_{t=1}^{T} D_{KL}\left(q_\phi\left(\mathbf{z}|\mathbf{x}^{(i)}\right) || \, p_\psi(\mathbf{z}_t)\right)$$

$\phi : Parameter\ of\ Noise\ Generator$

$\psi : Parameter\ of\ Prior\ Distribution$

$\mathbf{z} : Noise$

# <Learning Objective>

$$\mathcal{L}(\phi,\theta) = \lambda\mathcal{L}_{MLE}(\theta) + (1-\lambda)\mathcal{L}_{noise}(\phi,\theta)$$

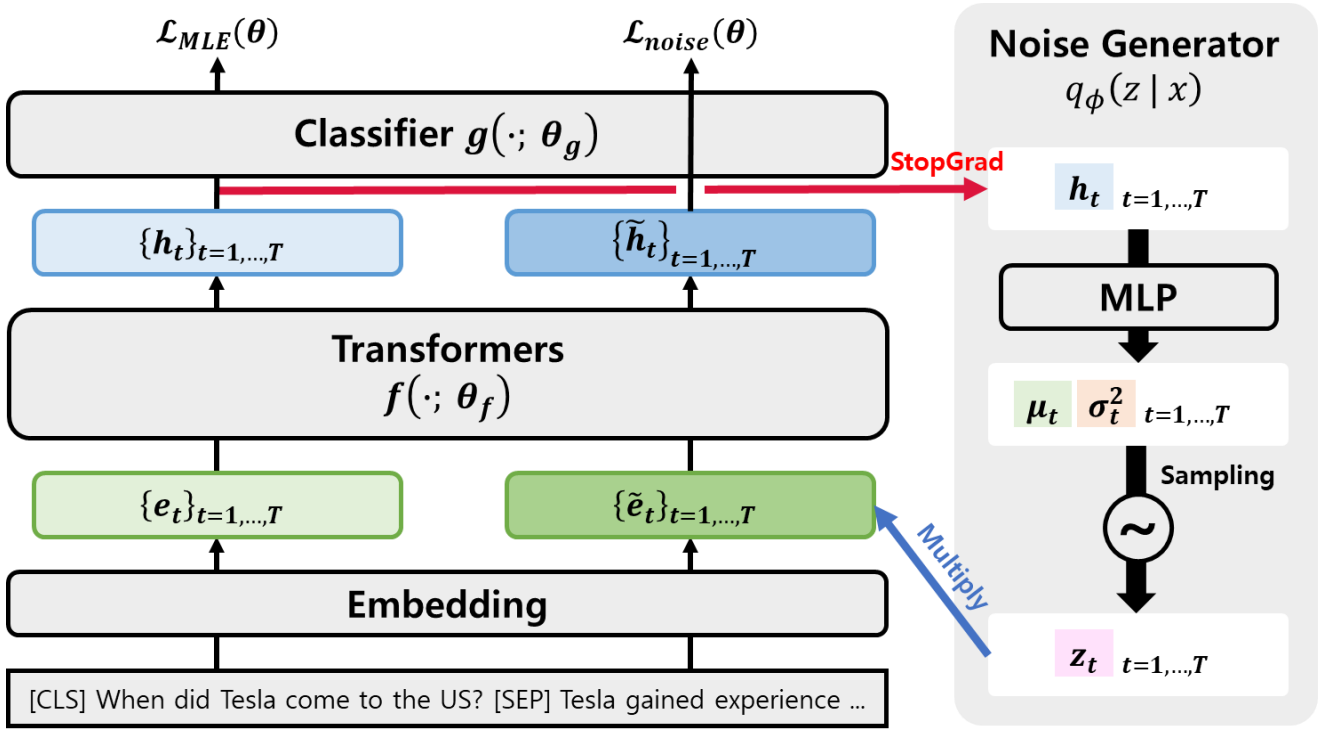
$$\mathcal{L}_{MLE}(\theta) := \sum_{i=1}^{N} \log p_\theta\left(\mathbf{y}^{(i)} \mid \mathbf{x}^{(i)}\right)$$

$$\mathcal{L}_{noise}(\phi,\theta) := \sum_{i=1}^{N} \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x}^{(i)})}\left[\log_{p_\theta}\left(\mathbf{y}^{(i)}|\mathbf{x}^{(i)},\mathbf{z}\right)\right] - \beta\sum_{t=1}^{T} D_{KL}\left(q_\phi\left(\mathbf{z}|\mathbf{x}^{(i)}\right) \,||\, p_\psi(\mathbf{z}_t)\right)$$

$\phi$ : *Parameter of Noise Generator*

$\psi$ : *Parameter of Prior Distribution*

$\mathbf{z}$ : *Noise*



# <Learning Objective>

$$\mathcal{L}(\phi,\theta) = \lambda\mathcal{L}_{MLE}(\theta) + (1-\lambda)\mathcal{L}_{noise}(\phi,\theta)$$

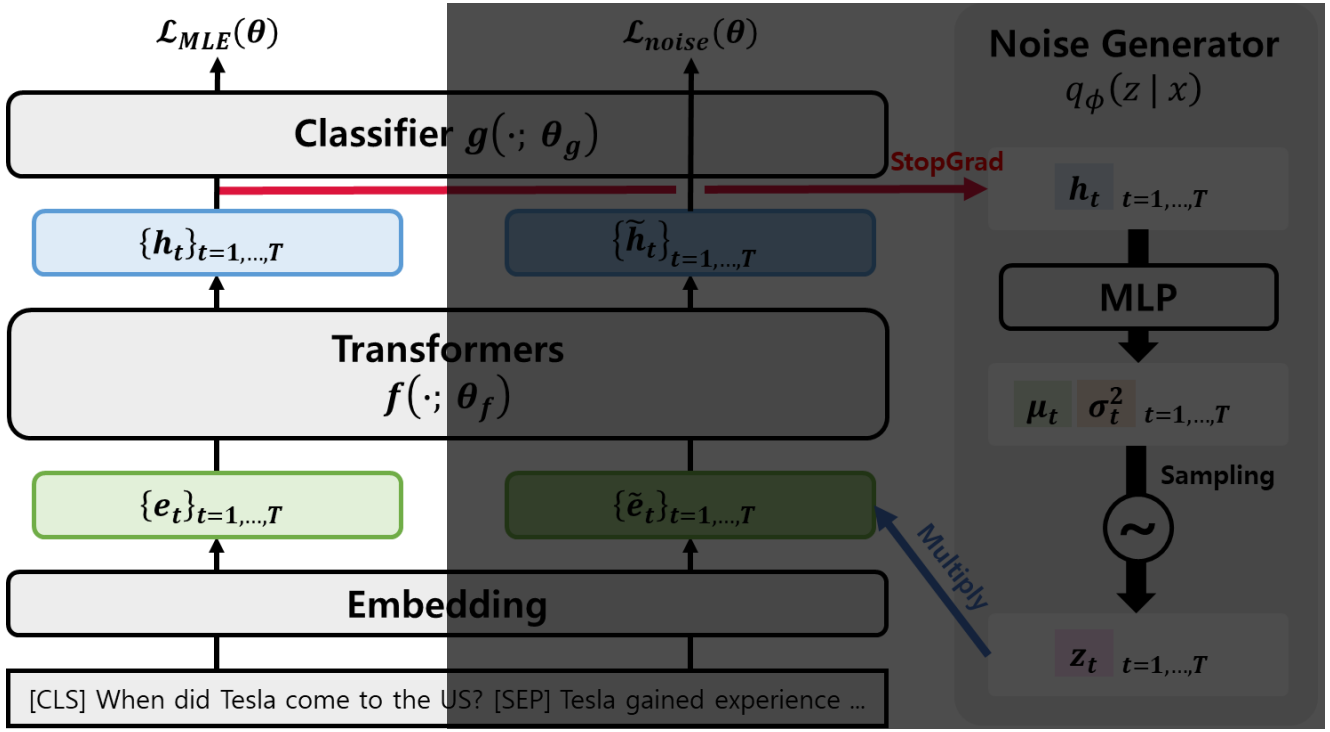
$$\boldsymbol{\mathcal{L}_{MLE}(\theta) := \sum_{i=1}^{N} log\, p_\theta\left(\mathbf{y}^{(i)} \mid \mathbf{x}^{(i)}\right)}$$

$$\mathcal{L}_{noise}(\phi,\theta) := \sum_{i=1}^{N} \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x}^{(i)})}\left[\log_{p_\theta}\left(\mathbf{y}^{(i)}|\mathbf{x}^{(i)},\mathbf{z}\right)\right] - \beta \sum_{t=1}^{T} D_{KL}\left(q_\phi\left(\mathbf{z}|\mathbf{x}^{(i)}\right) \,||\, p_\psi(\mathbf{z}_t)\right)$$

$\phi$ : *Parameter of Noise Generator*

$\psi$ : *Parameter of Prior Distribution*

$\mathbf{z}$ : *Noise*



# Method

## -Learning Objective

## <Learning Objective>

$$\mathcal{L}(\phi,\theta) = \lambda\mathcal{L}_{MLE}(\theta) + (1-\lambda)\mathcal{L}_{noise}(\phi,\theta)$$

$$\mathcal{L}_{MLE}(\theta) := \sum_{i=1}^{N} log\, p_\theta\left(\mathbf{y}^{(i)} \mid \mathbf{x}^{(i)}\right)$$

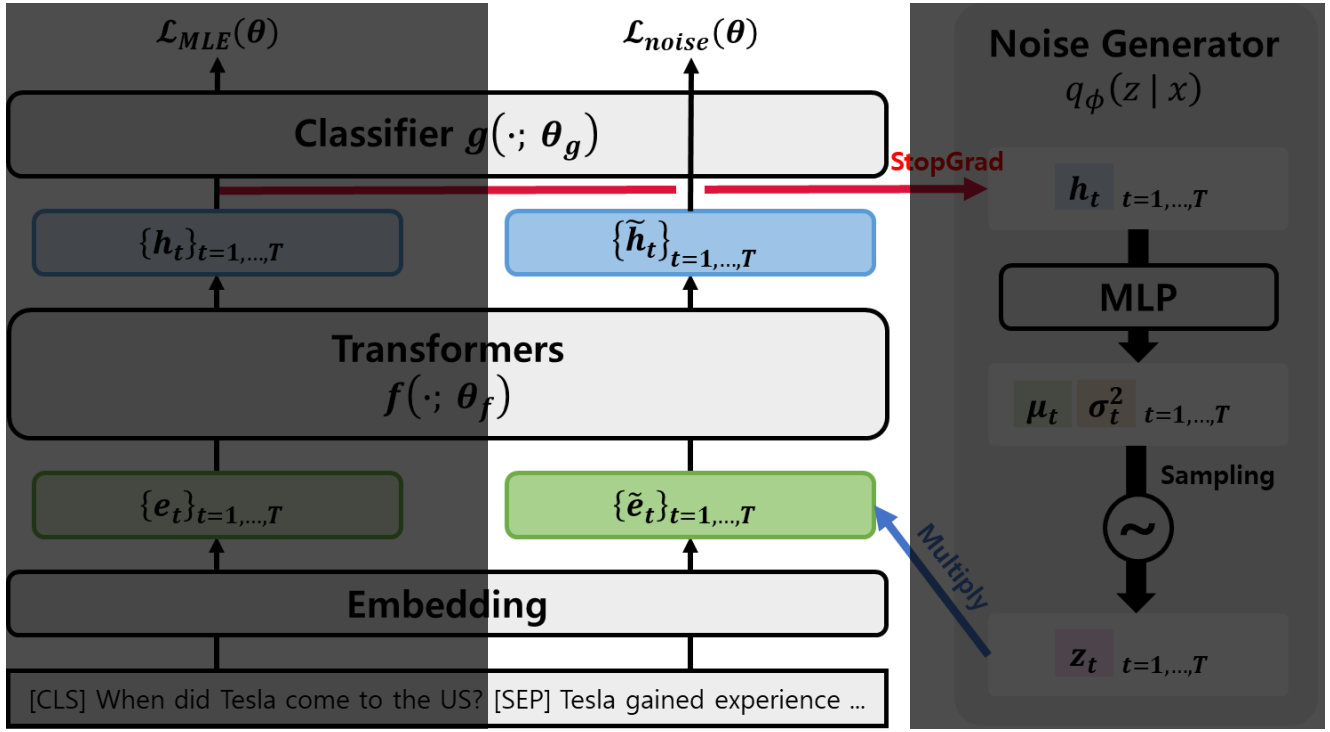
$$\boldsymbol{\mathcal{L}_{noise}(\phi,\theta) := \sum_{i=1}^{N} \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x}^{(i)})}\left[\log p_\theta\left(\mathbf{y}^{(i)}|\mathbf{x}^{(i)},\mathbf{z}\right)\right]}$$

$$-\beta \sum_{t=1}^{T} D_{KL}\left(q_\phi\left(\mathbf{z}|\mathbf{x}^{(i)}\right) \,||\, p_\psi(\mathbf{z}_t)\right)$$

$\phi$ : *Parameter of Noise Generator*

$\psi$ : *Parameter of Prior Distribution*

$\mathbf{z}$ : *Noise*



# <Learning Objective>

$$\mathcal{L}(\phi, \theta) = \lambda \mathcal{L}_{MLE}(\theta) + (1 - \lambda)\mathcal{L}_{noise}(\phi, \theta)$$

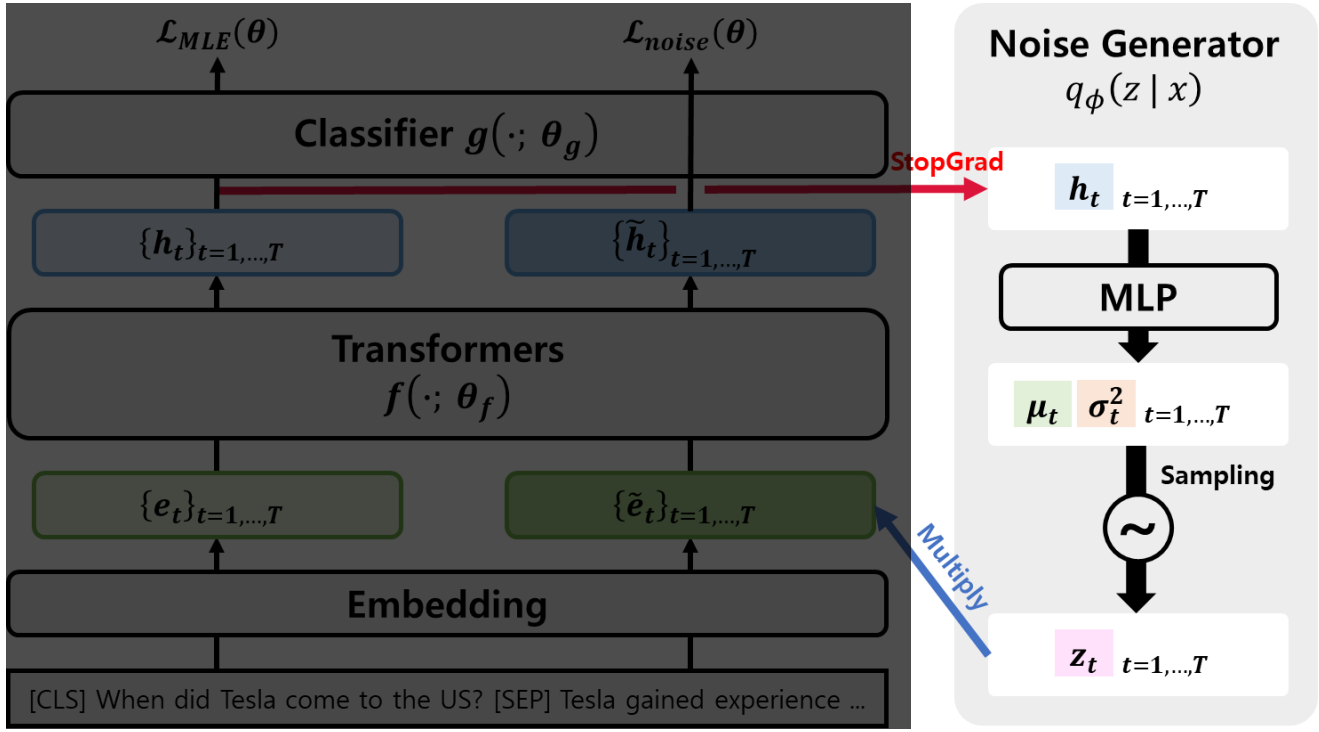
$$\mathcal{L}_{MLE}(\theta) := \sum_{i=1}^{N} log\, p_\theta \left(\mathbf{y}^{(i)} \mid \mathbf{x}^{(i)}\right)$$

$$\mathcal{L}_{noise}(\phi, \theta) := \sum_{i=1}^{N} \mathbb{E}_{q_\phi(\mathrm{z}|\mathrm{x}^{(i)})}[\log_{p_\theta}(\mathrm{y}^{(i)}|\mathrm{x}^{(i)}, \mathrm{z})]$$
$$-\beta \sum_{t=1}^{T} D_{KL}\left(q_\phi(\mathbf{z}|\mathbf{x}^{(i)}) \,||\, p_\psi(\mathbf{z}_t)\right)$$

$\phi$ : *Parameter of Noise Generator*

$\psi$ : *Parameter of Prior Distribution*

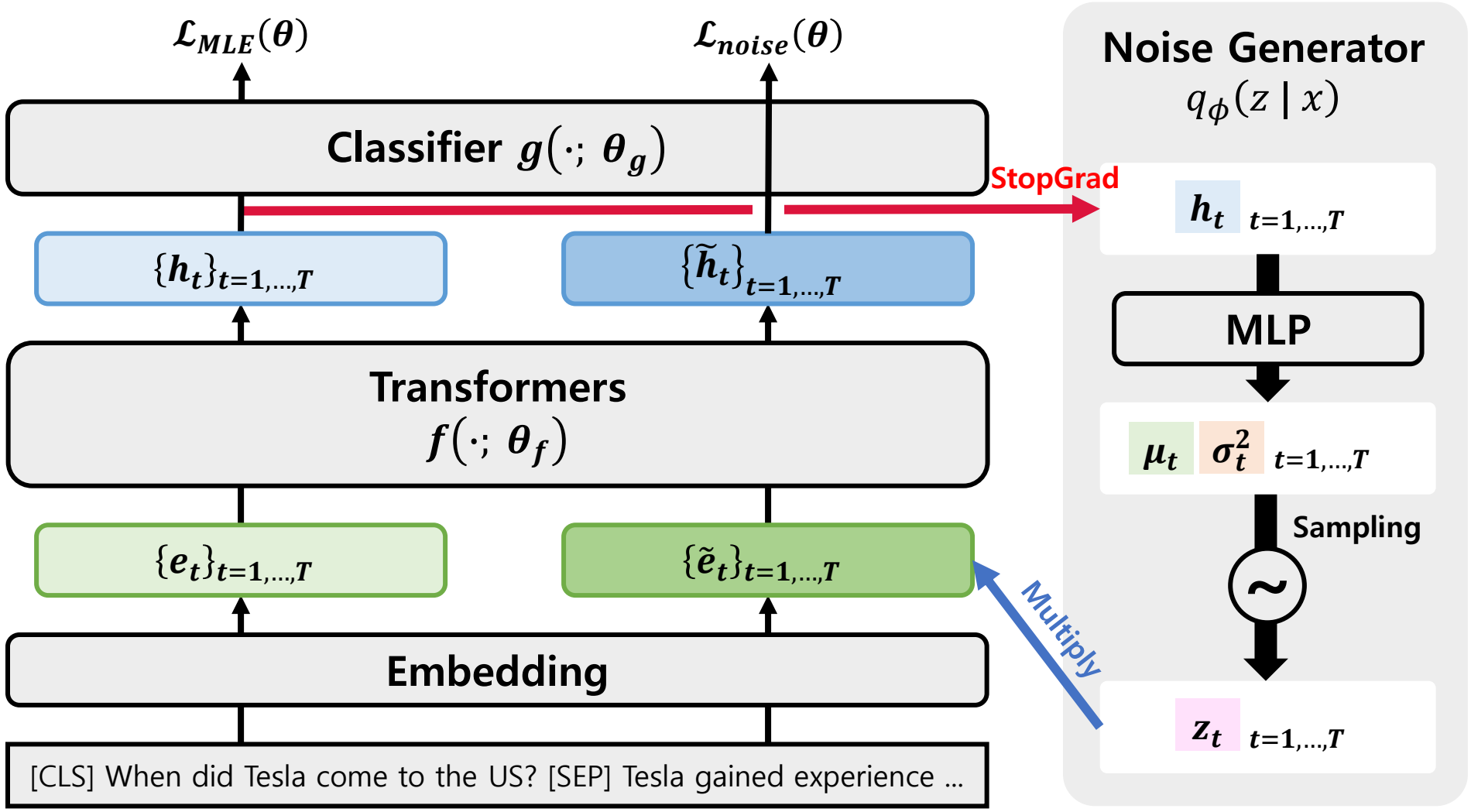
$\mathbf{z}$ : *Noise*



# Method

-Learning to Perturb Word Embedding

## <Overall Architecture>



# Experiments

- Experimental Results
- Quantitative Analysis
- Qualitative Analysis

# Experiments

- Experimental Results

## <Result>

| Method | SQuAD | Wiki | NYT | BioASQ | Reddit | Amazon |
|---|---|---|---|---|---|---|
| **BERT-base-uncased** (EM / F1) | | | | | | |
| **MLE** | 81.32 / 88.62 | 76.42 / 87.02 | 77.54 / 86.54 | 45.34 / 59.77 | 63.94 / 76.97 | 60.74 / 75.38 |
| **Adv-Aug** | 81.39 / 88.71 | 77.29 / 88.38 | 77.67 / 86.53 | 45.47 / 60.30 | 64.55 / 77.61 | 61.38 / 75.83 |
| **Word-Dropout** | 81.03 / 88.21 | 76.94 / 87.30 | 76.67 / 85.99 | 44.34 / 58.93 | 65.05 / 77.96 | 60.87 / 75.71 |
| **Gaussian-Dropout** | 81.47 / 88.78 | 77.28 / 87.23 | 77.25 / 86.35 | 45.27 / 61.37 | 65.19 / 77.73 | 61.67 / 75.98 |
| **Bernoulli-Dropout** | 81.46 / 88.76 | 77.34 / 87.40 | 77.16 / 86.35 | 44.21 / 59.33 | 64.53 / 77.25 | 61.27 / 75.85 |
| **SSMBA** | 78.17 / 86.53 | 74.33 / 85.26 | 74.31 / 83.98 | 39.96 / 54.49 | 59.29 / 73.50 | 56.57 / 71.81 |
| **Prior-Aug** | 81.77 / 89.04 | 77.95 / 87.83 | 77.92 / 86.81 | 46.40 / 60.80 | 65.50 / 78.16 | 61.57 / 76.22 |
| **SWEP** | **82.24 / 89.43** | **78.60 / 88.28** | **78.11 / 86.92** | **47.27 / 61.72** | **65.93 / 78.45** | **62.42 / 76.84** |
| **ELECTRA-small-uncased** (EM / F1) | | | | | | |
| **MLE** | 76.95 / 84.92 | 73.57 / 84.30 | 73.68 / 82.93 | 38.63 / 54.32 | 59.59 / 72.33 | 57.93 / 72.06 |
| **Adv-Aug** | 75.81 / 84.40 | 73.69 / 84.23 | 73.37 / 82.89 | 38.23 / 53.4 | **59.97** / 73.33 | 59.44 / 73.36 |
| **Word-Dropout** | 75.81 / 84.19 | 72.94 / 83.90 | 72.96 / 82.24 | 39.29 / 54.02 | 59.04 / 72.12 | 58.49 / 72.41 |
| **Gaussian-Dropout** | 76.42 / 84.53 | 73.31 / 84.11 | 73.27 / 82.51 | 37.30 / 52.46 | 59.29 / 72.31 | 57.50 / 71.65 |
| **Bernoulli-Dropout** | 76.31 / 84.50 | 73.50 / 84.08 | 73.35 / 82.75 | 37.10 / 52.37 | 59.33 / 72.56 | 57.71 / 71.99 |
| **SSMBA** | 77.75 / 85.81 | **74.90 / 85.21** | 73.25 / 82.62 | 39.02 / 53.32 | 58.97 / 72.83 | 56.66 / 71.89 |
| **Prior-Aug** | 77.70 / 85.60 | 74.65 / 85.02 | 74.38 / 83.47 | 38.96 / 54.19 | 59.92 / 73.10 | 59.01 / 73.11 |
| **SWEP** | **77.78 / 85.86** | 74.25 / 85.20 | **75.18 / 84.18** | **40.35 / 55.72** | 59.68 / **73.97** | **60.89 / 74.06** |

**Experimental results of extractive QA with BERT and ELECTERA model on six different test dataset**

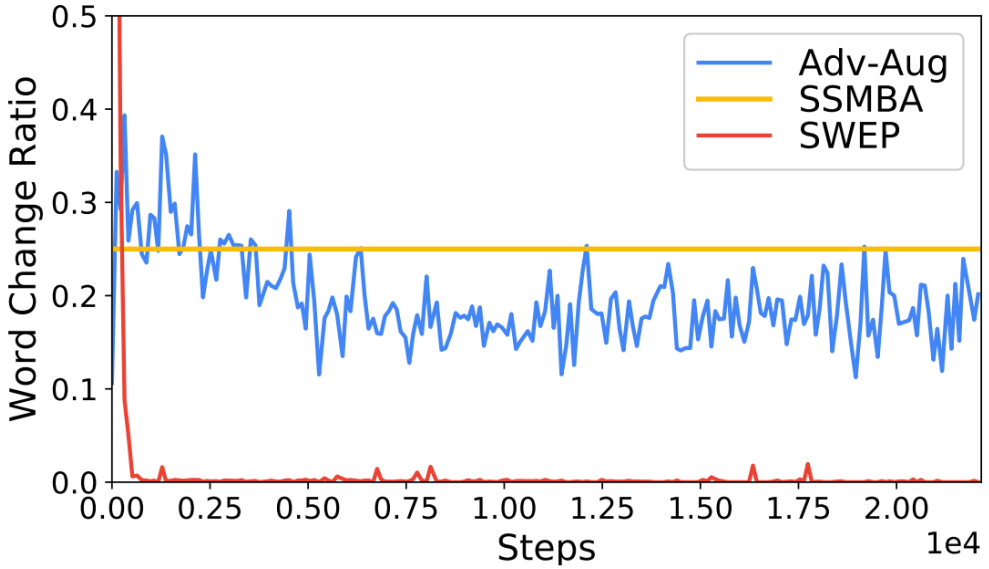
# <Result>

| Method | SQuAD | Wiki | NYT | BioASQ | Reddit | Amazon |
|---|---|---|---|---|---|---|
| | | | **T5-small** (EM / F1) | | | |
| **MLE** | **77.19** / 85.66 | 72.88 / 84.17 | 75.10 / 83.88 | 40.82 / 54.18 | **61.19 / 74.25** | 57.52 / 72.16 |
| **Adv-Aug** | 74.90 / 84.19 | 71.03 / 82.94 | 73.46 / 82.84 | 38.76 / 52.79 | 58.78 / 72.57 | 54.73 / 70.10 |
| **Word-Dropout** | 75.20 / 84.33 | 72.19 / 83.46 | 74.27 / 83.24 | 38.96 / 52.84 | 59.32 / 72.40 | 55.58 / 70.49 |
| **Gaussian-Dropout** | 76.25 / 84.86 | 72.56 / 83.69 | 74.76 / 83.57 | 41.15 / 54.64 | 60.14 / 73.40 | 57.01 / 71.52 |
| **Bernoulli-Dropout** | 75.15 / 84.34 | 71.64 / 83.33 | 73.81 / 83.06 | 39.42 / 53.77 | 59.06 / 72.48 | 55.22 / 70.46 |
| **SSMBA** | 74.94 / 84.19 | 71.97 / 83.85 | 73.29 / 82.79 | 37.96 / 51.57 | 58.54 / 72.51 | 55.05 / 70.62 |
| **Prior-Aug** | 76.88 / 85.47 | 73.11 / 84.18 | 75.52 / 84.04 | 40.49 / 54.47 | 60.92 / 74.04 | **57.99 / 72.38** |
| **SWEP** | 77.12 / **85.67** | **73.34 / 84.35** | **76.42 / 84.81** | **43.01 / 55.80** | 60.78 / 73.93 | 57.75 / 72.20 |

**Experimental results of generative QA with T5-small model on six different test dataset**

## Experiments
- Quantitative Analysis

# <Quantitative Analysis>




**Plot the extent to how many words changed by perturbation during training**

$$(v_1, \dots, v_d)^\top = \mathbf{W}_e^\top \tilde{\mathbf{e}}_t$$

$$j = \arg\max_i \{v_1, \dots, v_i, \dots, v_d\}$$

$$\tilde{\mathbf{w}}_t = \text{one} - \text{hot}(j, |V|)$$

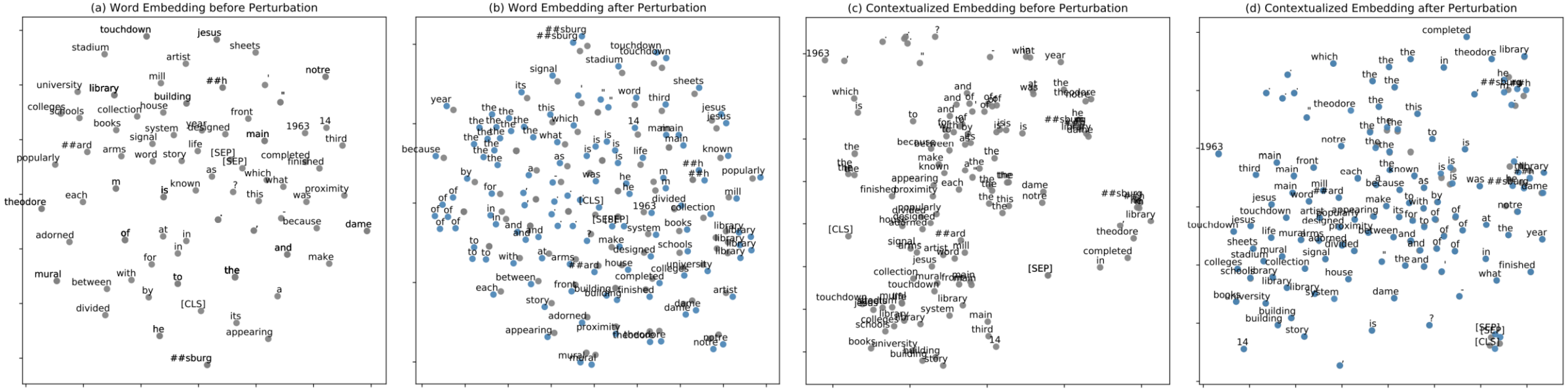
$\mathbf{W}_e$ : *Embedding Matrix*

$\tilde{\mathbf{e}}_t$ : *Perturbed Embedding*

# Experiments

- Qualitative Analysis

## <Qualitative Analysis>




**Overview of how the input is perturbed with SWEP.**

Contextualized embedding indicates the hidden states from the last layer of transformers.

# Conclusion

# <Conclusion>

- Proposed a simple yet effective data augmentation method to improve the generalization performance of pretrained language models for QA
- Showed that learned input-dependent perturbation function transforms the original input without changing its semantics
- Validated method for domain generalization tasks on diverse datasets, on which it largely outperforms strong baselines

# Any Questions?

# Thank You