

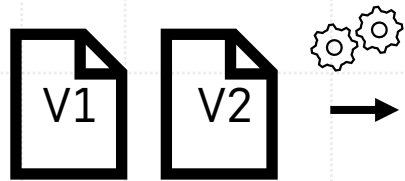


Clasificación automática de diferencias entre versiones de SLA

Juan Carlos Cortés Muñoz

María Elena Molino Peña

Descripción del problema planteado



Instance-Level SLA

NEW

V2

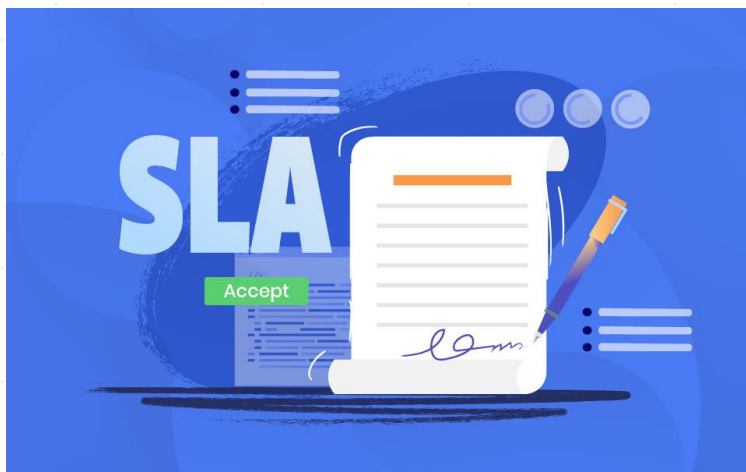
For each individual Amazon EC2 instance (“Single EC2 Instance”), AWS will use commercially reasonable efforts to make the Single EC2 Instance available with an **Instance-Level Uptime Percentage of at least 99.5%**, in each case during any monthly billing cycle (the “Instance-Level SLA”).

- Servicios
- **Métricas**
- **Objetivos**
- Compensaciones
- Reclamaciones
- Excepciones
- Definiciones

Conjunto de datos

▪ SLA Dataset

Oraciones de Acuerdos a nivel de Servicios **clasificadas** según su tipo



<https://github.com/marmolpen3/mle-sla-classification/blob/main/train.csv>

Variables:

- Text
- Service
- Metric
- Objective
- Remedy
- Claim
- Exception
- Definition
- Obligation
- Right
- Neither

Objetivo:

clasificar nueva oración en las categorías adecuadas

Algoritmos Transformers



Técnicas y herramientas

- Extracción, clasificación y **Dataset**
- **Zero** – Shot
- **Ajuste** de modelo Bert
- **Few** – Shot
- Comparativa y **predicción**
- **Publicación** en Hugging Face

Extracción, clasificación y Dataset

1 Selección de Acuerdos (!=proveedores)

3 Análisis de las oraciones

2 División del texto en oraciones (spacy)

4 Etiquetado de oraciones

	text	service	metric	objective	remedy	claim	exception	definition	obligation	right	neither
0	Last Updated: May 25, 2022 This Amazon Compute...	1	1	0	0	0	0	0	0	0	1
1	In the event of a conflict between the terms o...	1	0	0	0	0	0	1	0	0	1
2	Capitalized terms used herein but not defined ...	1	0	0	0	0	0	0	0	0	1
3	*For purposes of this SLA, Amazon EC2 includes...	1	1	0	0	0	0	0	0	0	1
4	SLAs AWS makes two SLA commitments for Amazon ...	1	1	0	0	0	0	0	0	0	1

Ajuste de modelo Bert

Evaluación en cada época

Epoch	Training Loss	Validation Loss	F1	Roc Auc	Accuracy
1	No log	0.499385	0.000000	0.500000	0.000000
2	No log	0.440801	0.000000	0.500000	0.000000
3	No log	0.376052	0.444444	0.646194	0.196078
4	No log	0.343754	0.649573	0.760381	0.470588
5	No log	0.327435	0.655462	0.766003	0.529412

Parámetros utilizados

Épocas	20
Cuota de aprendizaje	2e-5
Tamaño lote	4

Few – Shot

Entrenamiento de modelos con pocos ejemplos con el framework SetFit



Modelos utilizados:

all-**MiniLM**-L6-v2

paraphrase-MiniLM-L3-v2

paraphrase-**mpnet**-base-v2

Para clasificación multi etiqueta: multi_target_strategy="**one-vs-rest**"

Comparativa y predicción


Modelos	Rendimiento (Accuracy)
all-MiniLM-L6-v2	0.5294
bart-large-mnli	0.2549
bert	0.5294
paraphrase-MiniLM-L3-v2	0.5490
paraphrase-mpnet-base-v2	0.4313

Parámetros utilizados


Épocas	3
Cuota de aprendizaje	2e-5
Tamaño lote	12
Iteraciones por época	50
Función de pérdida	CSL

Publicación en Hugging Face


Models 3

 marmolpen3/all_MiniLM_L6_v2-sla

 • Updated about 4 hours ago

 marmolpen3/paraphrase-MiniLM-L3-v2-sla

 • Updated about 5 hours ago

 marmolpen3/bert-finetuned-sla

 • Updated about 5 hours ago • ↓ 4

⚡ Hosted inference API ⓘ

 Text Classification

Examples ▾

For each individual Amazon EC2 instance (“Single EC2 Instance”), AWS will use commercially reasonable efforts to make the Single EC2 Instance available with an Instance-Level Uptime Percentage of at least 99.5%, in each case during any monthly billing cycle (the “Instance-Level SLA”).

Compute

Computation time on Intel Xeon 3rd Gen Scalable cpu: 0.085 s

service	0.482
claim	0.361
metric	0.354
objective	0.286
definition	0.277



Conclusiones

Conjunto de datos está muy **limitado**

- Falta de datos sobre el dominio
- Clasificación manual

Limitaciones en el **hardware**

- Coste computacional al entrenar
- Tiempo de espera

Aprendizaje sobre **Hugging Face**

- Base para continuar investigando
- Alternativas NPL



Muchas gracias por su atención

Juan Carlos Cortés Muñoz
María Elena Molino Peña