



Becas colaboración curso 2020/2021

Fecha: 19 Junio 2020

Vicerrectorado de Investigación, Innovación y Transferencia

Subcomisión de I+D+i

Propuesta del departamento *SISTEMAS INFORMÁTICOS Y COMPUTACIÓN*

Núm Proyecto: 2020/32/00022

Responsable

Casacuberta Nolla, Francisco

E-mail

fcn@prhlt.upv.es

Ext.

73515

Responsable

Sánchez Peiró, Joan Andreu

E-mail

jandreu@prhlt.upv.es

Ext

77358

Título proyecto

Traducción de textos basada en redes neuronales.

Valoración proyecto

4

Descripción proyecto

Hoy en día, las traducciones de texto que se obtienen de forma automática han alcanzado un alto grado de calidad gracias a las redes neuronales y aprendizaje profundo a partir de conjuntos de aprendizaje, esto es, pares de frases y sus correspondientes traducciones. La técnica que está teniendo más éxito es el Transformer, un codificador-decodificador neuronal que explota al límite las relaciones entre sí de las palabras a traducir, las palabras traducidas y las relaciones entre las palabras a traducir y las traducidas. El objetivo de este trabajo es diseñar un traductor basado en esta tecnología y a partir de un corpus de aprendizaje del Europarl.

Actividades a realizar por el alumno

1. Estudiar el estado del arte de las arquitecturas encoder-decoder basadas en redes neuronales, y concretamente en Transformer, para la traducción de textos.
2. Familiarización con los paquetes de herramientas existentes sobre redes neuronales para la traducción.
3. Preparación de un corpus de entrenamiento a partir del Europarl.
4. Construcción de un prototipo para la traducción de textos basado en el Transformer utilizando el paquete de herramientas escogido.
5. Evaluación del prototipo construido mediante el análisis de las traducciones de un conjunto de test dado.
6. Redacción de la memoria donde se describa el trabajo realizado, así como las conclusiones que se puedan extraer y el trabajo futuro.

Horario

El horario se consensuará con el profesor para adaptarlo al horario de clases de el/la alumno/a.