

# Deep Learning techniques for advanced language understanding

PhD Candidate:

Moreno La Quatra

#### 1. Introduction

Massive amount of text data is published daily over online and offline platforms (e.g., Social Networks and News Editors). **Automatic Text Summarization** aims at analyzing an arbitrary amount of text to

#### 3. Multilingual adaptation [2]

We proposed a new methodology aimed at adapting general-purpose word embedding models of a target language (low-resource) to its domain-specific version.

 • •		• •

outline the most relevant content.

Deep neural networks can model human language. Leveraging state-of-the-art deep learning methodologies my research focuses on the analysis of multimedia content to generate effective data representations to address downstream tasks (e.g., text summarization).

## 2. Timeline Summarization [1]

TimeLine Summarization (TLS) aims at extracting timestamped summaries of longlasting news topics. We proposed a novel approach consisting of the following steps:

- **Date Summarization:** rank sentences according to their estimated relevance.
- **Graph Modelling:** combine both contentrelated and reference-based connections. Graph ranking is used to select relevant



Fig. 2: 4-step process for multilingual domain adaptation of word embedding models

## 4. Multimodal Learning

User-generated content is intrinsically multimodal. The goal of multimodal learning is to design architectures that can simultaneously analyze several modalities to benefit of multimedia data richness.



# dates.



**Fig. 1: Timeline Summarization Pipeline** 

Fig. 3: Single vs Multimodal learning.

[1] **La Quatra, Moreno,** et al. "Summarize Dates First: A Paradigm Shift in Timeline Summarization." Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval. 2021.

[2] Cagliero, Luca, and **Moreno La Quatra**. "Inferring multilingual dmain-specific word embeddings from large document corpora." IEEE Access (2021).

Complete publication list: <u>morenolaquatra.github.io/publications</u>