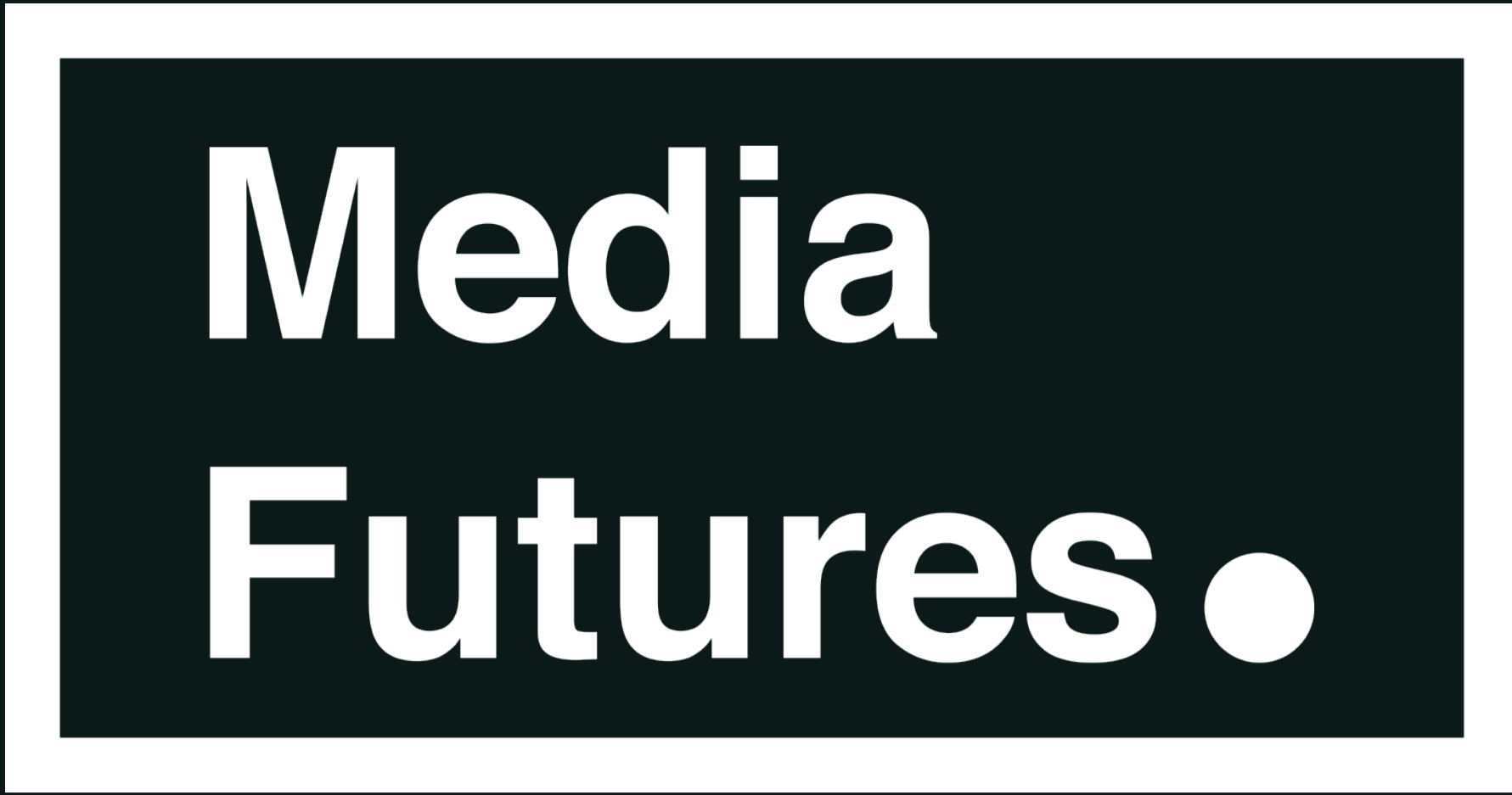# Editorial Component in a News Recommender System

## Media Futures.

Bilal Mahmood (WP2) I Bilal.Mahmood@uib.no

Supervisors: Mehdi Elahi (UiB), Samia Touileb (UiB), Lubos Steskal (Tv2), Lars Skjærven (Tv2)

## Introduction

**The objective of this project is to develop an impactful News Recommender System (NRS).** There are several ways in which an impactful news recommender system could be built. One important way is by **automating** how **editors recommend news articles** and add it along with **personalization**. Given a particular article, which articles an editor would pick, and accurately modeling it as a component in a NRS would add both personalization and editorial trust. To achieve that, modeling news articles is a crucial step and this work shows the research in that direction.

## Research question

**RQ1.** How to utilize advanced natural language processing (NLP) techniques to model news articles that can assist in the process of automating editors' recommendation of news articles?

## Methodology

**Embedding models** such as Text-Embedding-Ada-002, Sentence-BERT, NorBERT3, and similar counterparts, will be systematically investigated to capture the **semantic essence** of news articles. Multiple components such as title, lead title, body, media content, and other relevant features will be taken into account during this exploration. Subsequently, these semantically rich representations will undergo a rigorous evaluation process to gauge their efficacy in enhancing the precision of understanding how editors recommend news articles. With this understanding, **machine learning** (ML) models would be explored to automate the selection decisions that the editors make. These ML models would later be used in building a hybrid NRS that would comprise **editorial** as well as user **personalization** components.
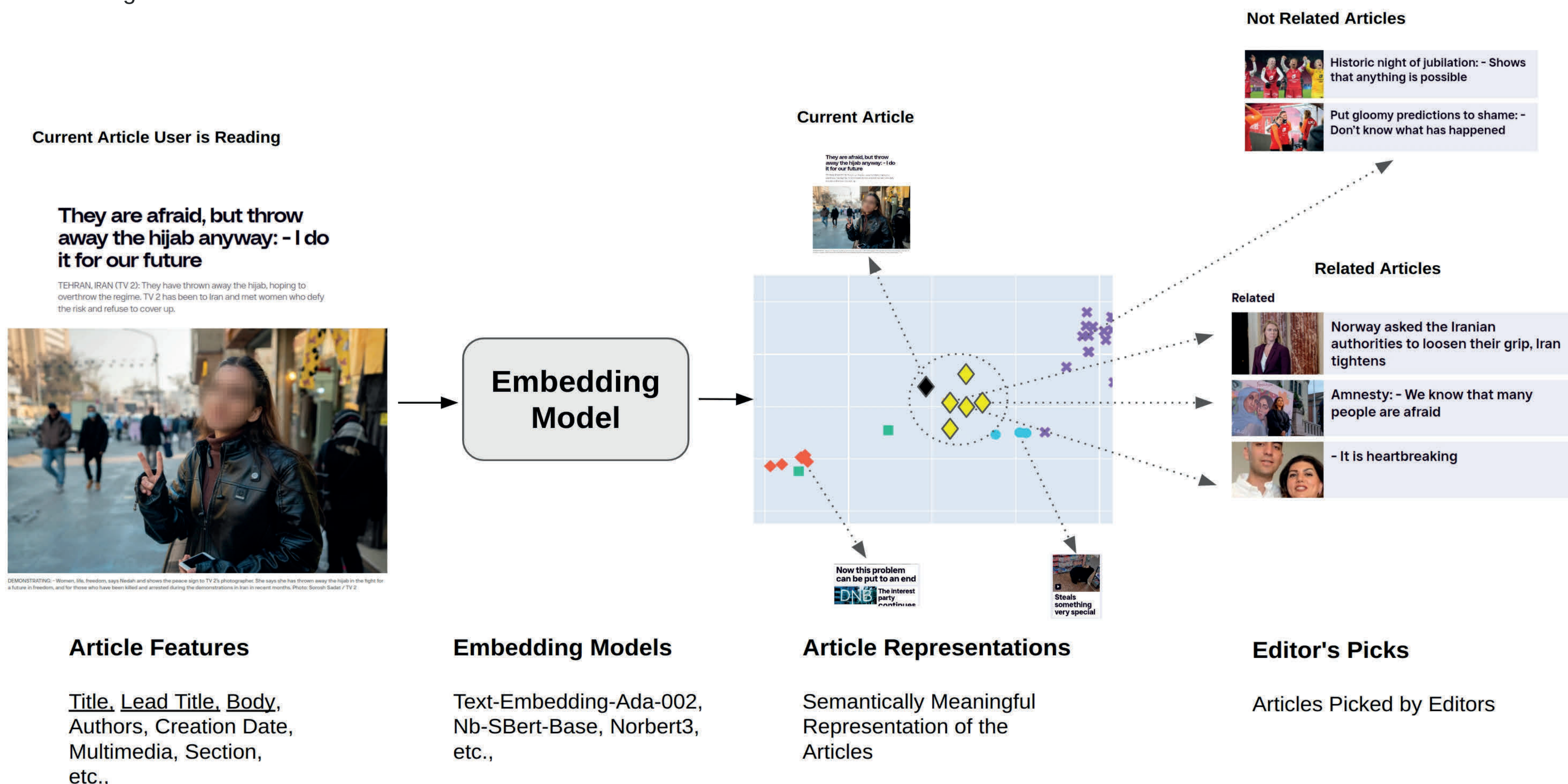


Figure 1: Modeling News Articles Using Latest Embedding Models for Automating Editorial Picks

**Article Features**

Title, Lead Title, Body, Authors, Creation Date, Multimedia, Section, etc.,

**Embedding Models**

Text-Embedding-Ada-002, Nb-SBert-Base, Norbert3, etc.,

**Article Representations**

Semantically Meaningful Representation of the Articles

**Editor's Picks**

Articles Picked by Editors

## PARTNERS

amedia  Bergens Tidende  HIGHSOFT  IBM  NRK  Schibsted  2  vizrt  S  Universitetet i Oslo  NORCE  nb.no  WOLFTECH  Faktisk.

## HOST

UNIVERSITY OF BERGEN

## FUNDED BY

*This research is funded by SFI MediaFutures partners and the Research Council of Norway (grant number 309339).*

Forskningsrådet   Norwegian Centre for Research-based Innovation