**Politics Fake News Detector in LATAM (Latin America)**

**The Background**

Since the Cambridge Analytica scandal a pandora box has been opened around the world, bringing to light campaigns even involving our current Latinamerica leaders manipulating public opinion through social media to win an election. There is a common and simple pattern that includes platforms such as facebook and fake news, where the candidates are able to build a nefarious narrative for their own benefit. This fact is a growing concern for our democracies, as many of these practices have been widely spread across the region and more people are gaining access to the internet. Thus, it is a necessity to be able to advise the population, and for that we have to be able to quickly spot these plots on the net before the damage is irreversible.

**The Problem**

**English**

Once the capacity to somewhat detect irregularities in the news activity on the internet is developed, we might be able to counter the disinformation with the help of additional research. As we reduce the time spent in looking for those occurrences, more time can be used in validating the results and uncovering the truth; enabling researchers, journalists and organizations to help people make an informed decision whether the public opinion is true or not, so that they can identify on their own if someone is trying to manipulate them for a certain political benefit.

If  this matter isn’t tackled with enough urgency, we might see the rise of a new dark era in latin america politics, where many unscrupulous parties and people will manage to gain power and control the lives of many people. Therefore, the results of the project can provide support for both private and public companies on their future analysis and activities. Additionally, researchers and students could use the outcomes for their own research or use it for learning purposes.

**Español**

Una vez contemos con la capacidad de detectar irregularidades en las noticias por internet, seremos capaces de contrarrestar la desinformación con la ayuda de investigaciones adicionales. Mientras reducimos el tiempo invertido en identificar estos patrones, podemos dedicar más tiempo a validar los resultados y buscar las verdades ocultas; habilitando a los investigadores, periodistas y organizaciones a que  ayuden a la población a tomar decisiones informados de la veracidad de la opinión pública, y que estos puedan identificar si alguien está tratando de manipularlos para beneficio político.

Si este problema no se trata con urgencia, podríamos ver el resurgir de una era oscura en la política latina, donde muchos partidos y personajes inescrupulosos tomarán el poder y el control de la vida de las personas. Por lo tanto, los resultados del proyecto pueden ser de provecho para los análisis y actividades futuros de entidades tanto públicas como privadas. Además de que tanto estudiantes como investigadores pueden hacer uso de los entregables para sus propias investigaciones y/o aprendizaje.

**The Project Goals**

– To gather and clean datasets from different newspapers and new outlets in LATAM.

– To predict if there is a political affiliation in a certain topic on the news.

– To compare and determine if there is any irregularity between the information available respecting an specific news in the different news sources

– To understand and visualize the information patterns from the news by creating a visualization dashboard.

**The Learning Outcomes**

* 1. How to gather and clean text datasets from news for data modeling.
* 2. How to use data visualization tools for further app creation and data reporting.
* 3. How to create a classification model with NLP libraries.

**The Tasks & Timeline**

|  |  |  |  |
| --- | --- | --- | --- |
| **Week 1** | **Week 2** | **Week 3** | **Week 4** |
| – Data collection | – Data collection  – Data cleaning | – Topics Analysis  – Data cleaning | – Unsupervised Model Creation |

|  |  |  |  |
| --- | --- | --- | --- |
| **Week 5** | **Week 6** | **Week 7** | **Week 8** |
| – Division by Branches | – Political Party Classifier (If feasible)  – Map Visualization | – Streamlit App  – Report | – Streamlit App  – Deployment  – Report |

**Proposals of workshops topics for the challenge:**

*Select the workshops you would like us to organize for your project from listed down below and share your thoughts and needs during the kick-off meeting. If you would like us to organize a workshop on a topic not listed here, please mention it here and we will try to find a speaker for it.*

* Data gathering and cleaning for News/text
* Topic modeling NLP
* Streamlit – creating interactive visualizations