Detect Cyber Bullying

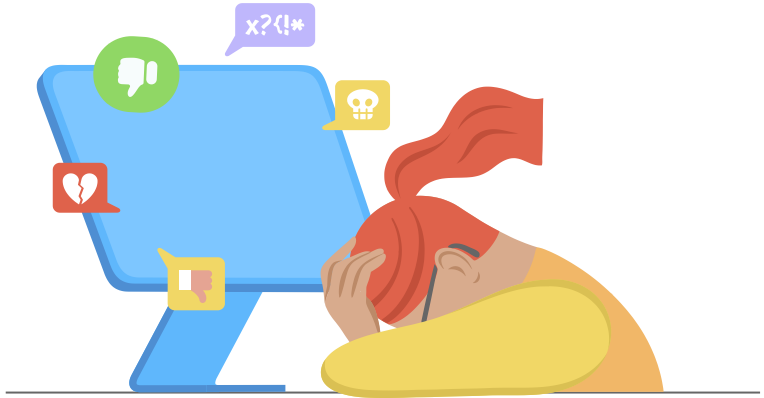
Final project

Ironhack Data Analytics Bootcamp 2023

Carmen Matos and Juliane Petersen

Trigger Warning

This presentation contains content about bullying.

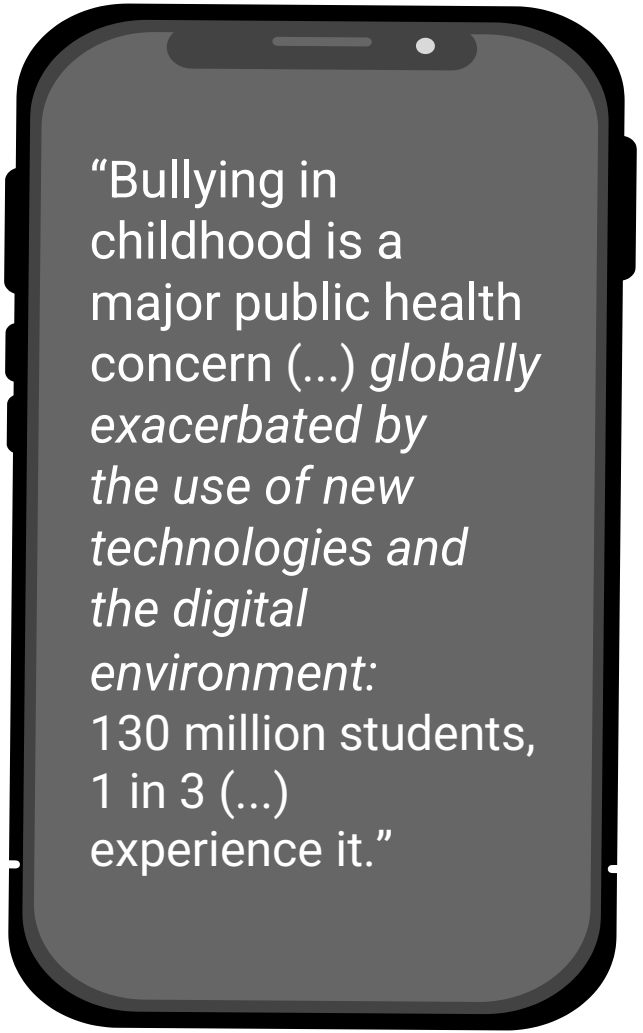


The problem:

United Nations | UNITED NATIONS HUMAN RIGHTS OFFICE OF THE HIGH COMMISSIONER

Panel discussion on cyberbullying against children

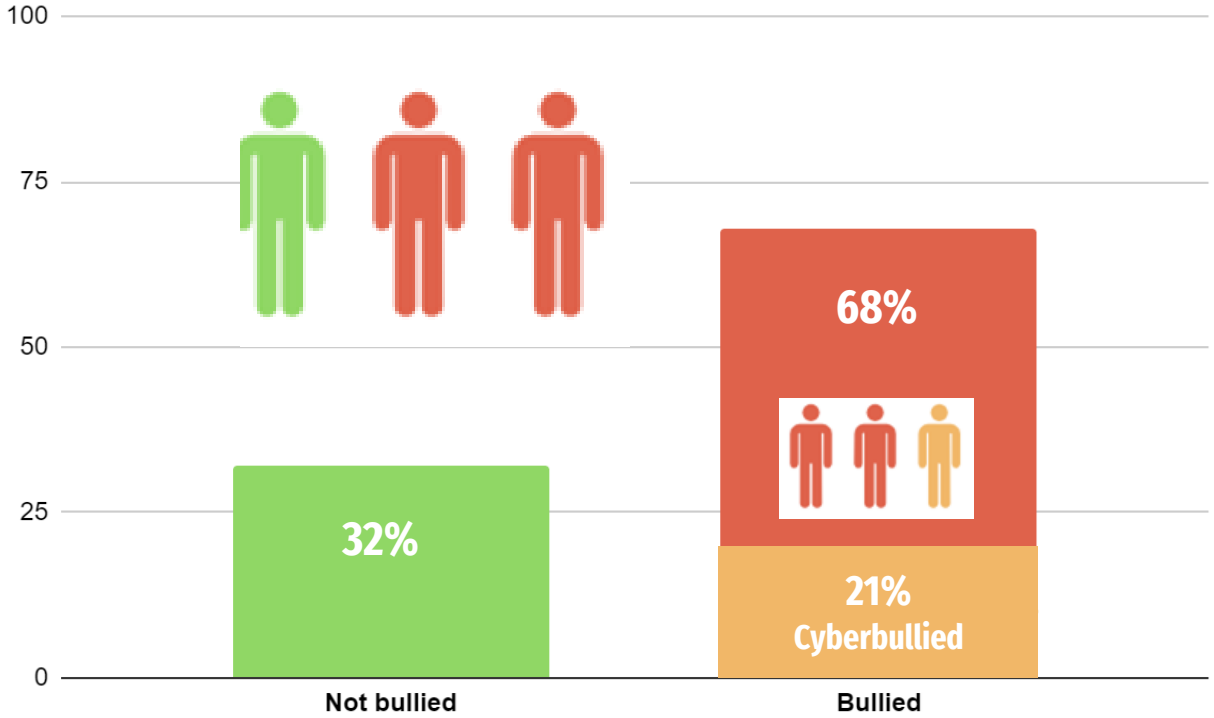
54th session of the UN Human Rights Council, September 2023



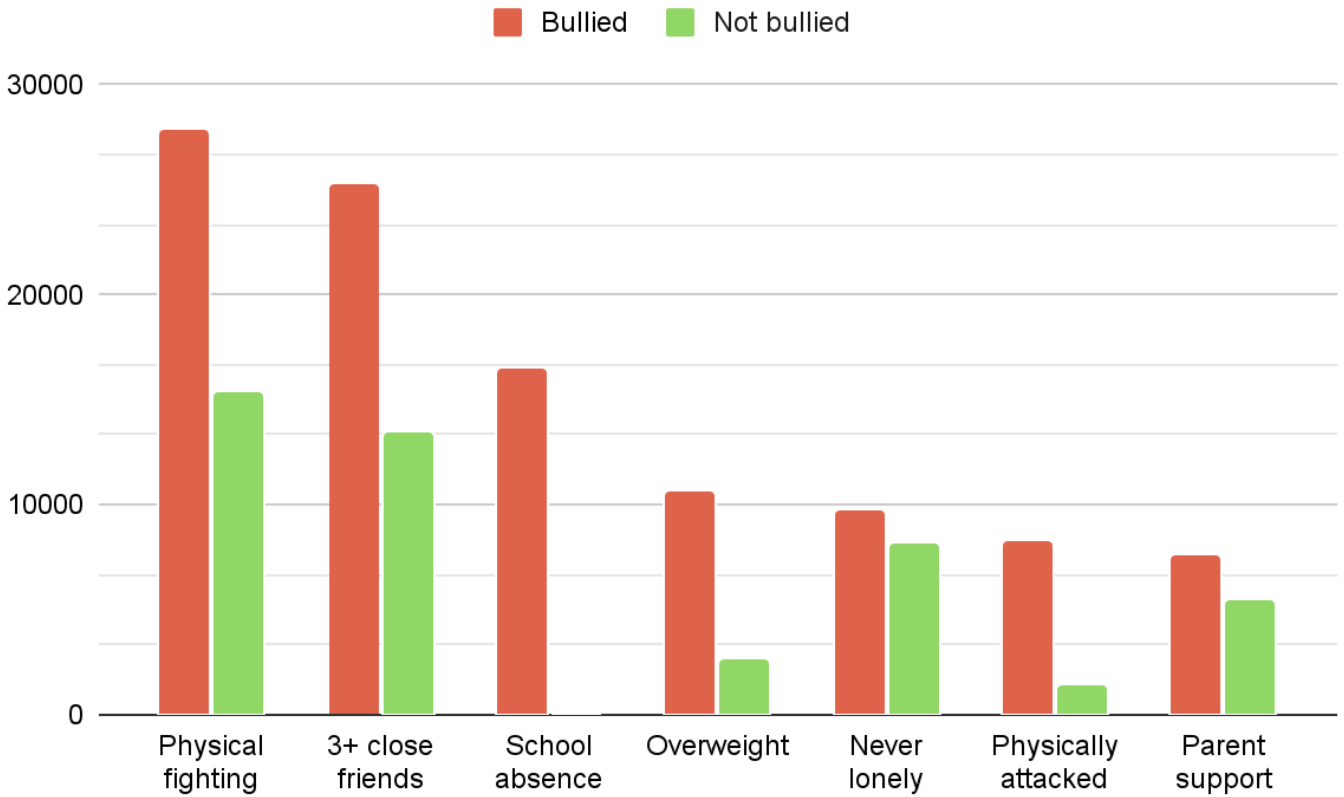
Goal 1: Identifying risk factors for bullying using data analysis

Dataset: "Bullying in schools" : 56,981 participants

(Global School-based Student Health Survey, Argentina, 2018, Source: Kaggle)



Main risk factors: Physical fighting, school absence, extra weight



Goal 2: Creating a tool to detect cyberbullying using a Machine Learning (ML) model and Natural Language Processing (NLP)

Targeting problems created by technology with even better technology

Dataset: "Cyberbullying Dataset"
159,388 messages from social media platforms classified as bullying or not (Source: Kaggle)

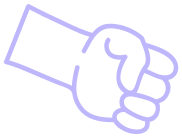
The process

01 **Text preprocessing (i.e. tokenizing)**

02 **Feature extraction (i.e. embedding)**

03 **Model selection, training & evaluation**

04 **Model testing & app Interface**



Challenges

Huge dataset == huge time consumption & reaching computational limits

New topic == trial & error, i.e. chunking, NO SMOTING, Multiple Classification, Streamlit

01 02

Data preprocessing: A simplified example

INPUT MESSAGE:

Thanks to all of you for the amazing time. We wish all of you a glorious future from the bottom of ours hearts. <3 Love, Carmen & Juli

Cleaned (“ “, !”, .lower, :-), ...)	
Stopwords & tokenization	
Lemmatization	
Embedding	

01 02

Data preprocessing: A simplified example

INPUT MESSAGE:

Thanks to all of you for the amazing time. We wish all of you a glorious future from the bottom of ours hearts. <3 Love, Carmen & Juli

Cleaned (“ “, !”, .lower, :-), ...)	thanks to all of you for the amazing time we wish all of you a glorious future from the bottom of ours hearts love carmen juli
Stopwords & tokenization	
Lemmatization	
Embedding	

01 02

Data preprocessing: A simplified example

INPUT MESSAGE:

Thanks to all of you for the amazing time. We wish all of you a glorious future from the bottom of ours hearts. <3 Love, Carmen & Juli

Cleaned (" ", !", .lower, :-), ...)	thanks to all of you for the amazing time we wish all of you a glorious future from the bottom of ours hearts love carmen juli
Stopwords & tokenization	thanks amazing time wish glorious future bottom hearts love carmen juli
Lemmatization	
Embedding	

01 02

Data preprocessing: A simplified example

INPUT MESSAGE:

Thanks to all of you for the amazing time. We wish all of you a glorious future from the bottom of ours hearts. <3 Love, Carmen & Juli

Cleaned (“ “, !”, .lower, :-), ...)	thanks to all of you for the amazing time we wish all of you a glorious future from the bottom of ours hearts love carmen juli
Stopwords & tokenization	thanks amazing time wish glorious future bottom hearts love carmen juli
Lemmatization	thank amazing time wish glorious future bottom heart love carmen juli
Embedding	

01 02

Data preprocessing: A simplified example

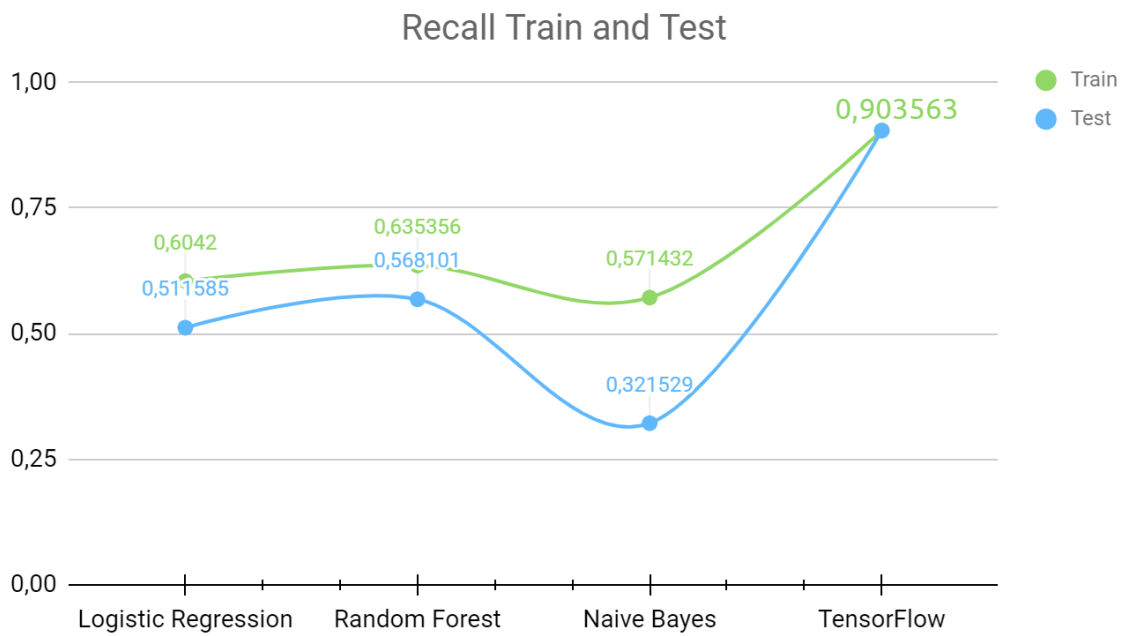
INPUT MESSAGE:

Thanks to all of you for the amazing time. We wish all of you a glorious future from the bottom of ours hearts. <3 Love, Carmen & Juli

Cleaned (“ “, !”, .lower, :-), ...)	thanks to all of you for the amazing time we wish all of you a glorious future from the bottom of ours hearts love carmen juli
Stopwords & tokenization	thanks amazing time wish glorious future bottom hearts love carmen juli
Lemmatization	thank amazing time wish glorious future bottom heart love carmen juli
Embedding	[[28324], [11508], [16234], [28324], [10119], [], [16234], [28324], [29233], [16234], [], [22326], [5330], [11508], [16234], [], [12887], [22641], [5330], [22326], [9754], [8001], ...

03

Model selection, training and evaluation



+90 %
Sequential Keras

+56 %
Random Forest

+51 %
Logistic Regression

+32 %
Naive Bayes

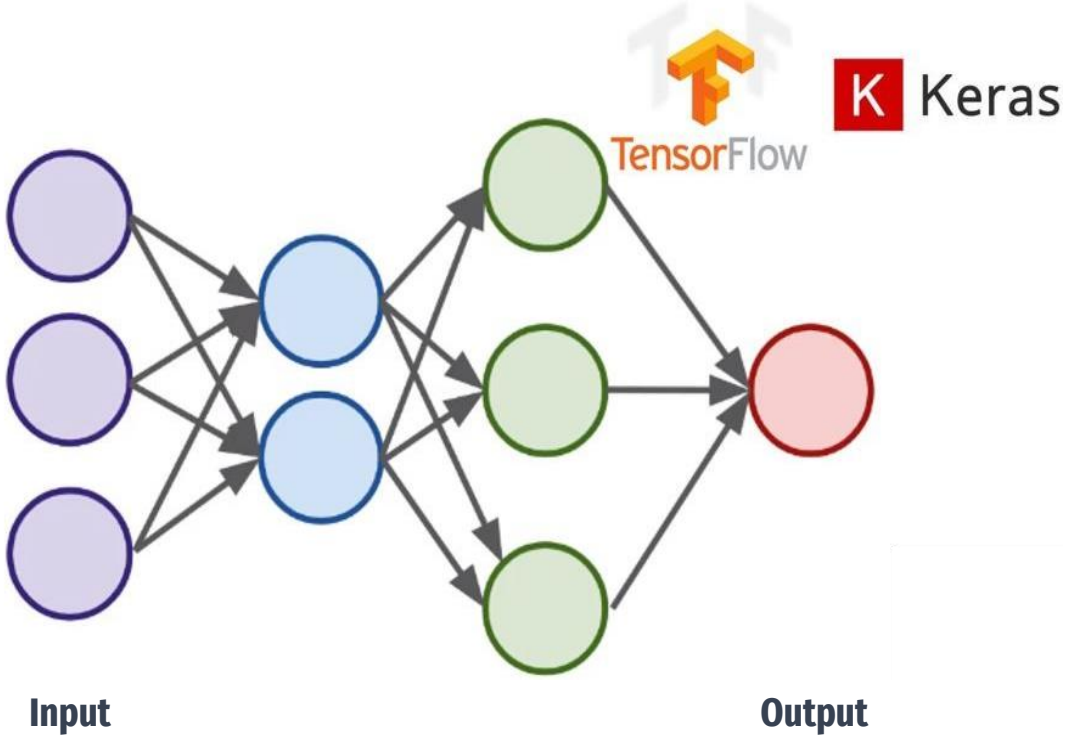
Sequential Model - Keras/TensorFlow

Pros:

- Powerful for complex patterns and relationships, i.e. NLP
- Learns hierarchical features from data

Cons:

- Requires more data and
- computational resources for training
- Careful tuning of hyperparameters



04 Model testing and app interface

Final Project

Data Analytics OCT/23

This is a Streamlit app for detecting cyberbullying content.

Carmen Matos and Juliane Petersen

Detect Cyberbullying

Enter a message:

Bullying or not?


Thank you!