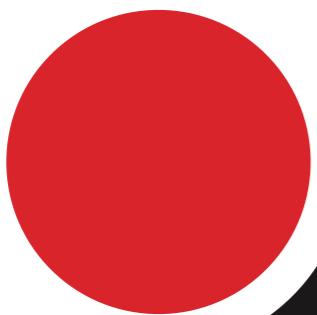


# How Do Fake-News Go Viral?

Or why Bernie Sanders could replace  
Trump with little-known loophole.



ÉCOLE POLYTECHNIQUE  
FÉDÉRALE DE LAUSANNE

EE-558: A Network Tour of Data Science

William Trouleau & Victor Kristof

# Could Sanders Replace Trump?

EDITION  
US

THE HUFFINGTON POST



NEWS POLITICS ENTERTAINMENT WELLNESS WHAT'S WORKING VOICES VIDEO ALL SECTIONS



TRENDING



Dan Rather Blasts White House Over 'Alternative Facts'

Trevor Noah Likes Seeing Kellyanne Conway 'Taste The Bulls\*\*t In Her Mouth'

Shia LeBeouf Expertly Shouts Down Alleged White Supremacist During Livestream

CONTRIBUTOR

## Bernie Sanders Could Replace President Trump With Little-Known Loophole

Read this article and then share with your friends.

11/14/2016 01:05 pm ET | Updated Nov 15, 2016



Matt Masur



Entrepreneur, New Media Personality & Progressive Activist



MATTHEW MASUR

Bernie Sanders in Syracuse, N.Y., 2016

Here is exactly what we need to do to save our great society. The information here is what we've all been waiting for. By doing this we can make Bernie the president on Inauguration day rather than President-elect [Donald Trump](#).

Actually, no we can't. There is no loophole that allows a random person to assume the office of president. That's pretty basic common sense but yet you clicked or even shared this article anyway. Now that right there is the real point of this post...

TRENDING

Fans Think This Photo Of Selena Gomez Is Photoshopped



Taylor Swift Tweets Support For Women's March But People Aren't Buying It



Trump 'Office' Parody Is A Glimpse At The Buffoonery We Have In Store



'Saturday Night Live' Writer Suspended For Tweet Mocking Barron Trump



Former Trump Staffer Attacks Cleveland Newspaper





# BuzzFeed fact checking

## BuzzFeedNEWS

News   Videos   Quizzes   Tasty   DIY   More ▾   [Get Our News App](#)

### Hyperpartisan Facebook Pages Are Publishing False And Misleading Information At An Alarming Rate

A BuzzFeed News analysis found that three big right-wing Facebook pages published false or misleading information 38% of the time during the period analyzed, and three large left-wing pages did so in nearly 20% of posts.

posted on Oct. 20, 2016, at 7:47 p.m.

**Craig Silverman**  
BuzzFeed Founding Editor,  
Canada

**Lauren Strapagiel**  
BuzzFeed News Reporter

**Hamza Shaban**  
BuzzFeed News Reporter

**Ellie Hall**  
BuzzFeed News Reporter

**Jeremy Singer-Vine**  
BuzzFeed News Reporter



BuzzFeedNews/2016-10-facebook-fact-check

# BuzzFeed fact checking

- Manually rated the **truthfulness** of all posts from **9** popular Facebook pages for one week in September 2016



# BuzzFeed fact checking

Right Wing News added a new photo.  
September 27, 2016 ·

**I USED BLACKS AND MINORITIES TO GET ELECTED AND DID NOTHING FOR THEM**

No factual content

**NOW I AM ENDORSING HILLARY CLINTON BECAUSE SHE IS BEST QUALIFIED TO DO NOTHING TOO**

<http://hillaryclintonsucks.xyz>

Like Comment Share

30K

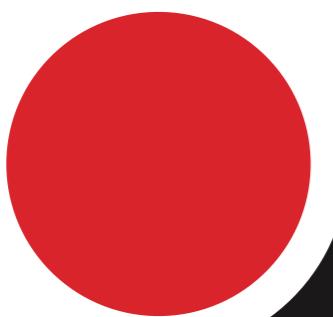
3,780 shares 37,632 shares

Top Comments 721 Comments

Comments

Can we use a **data-driven** approach  
to automatically **detect fake-news**  
**posts ?**





# Data Acquisition

# Data Acquisition Pipeline



For all posts in 2016:

- Post id
- Time of creation
- Type of post
- Text message of the post
- Number of shares
- Number of comments
- Number of reactions

Add BuzzFeed rating:

- Mostly true
- Mixture of true and false
- Mostly False
- No factual content

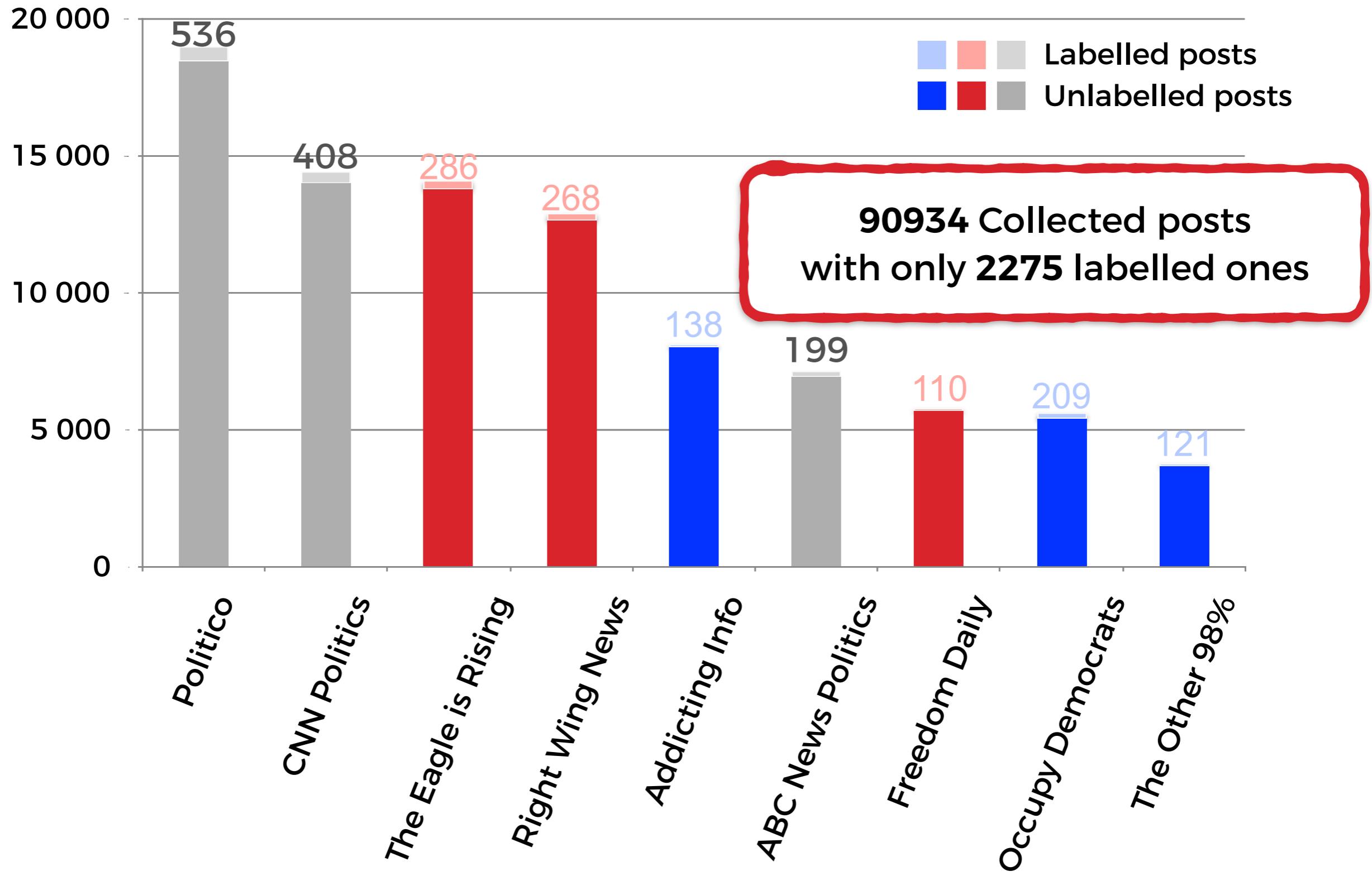
Add political orientation category:

- Pro-democrat page
- Pro-republican page
- Mainstream page



# Data Exploration

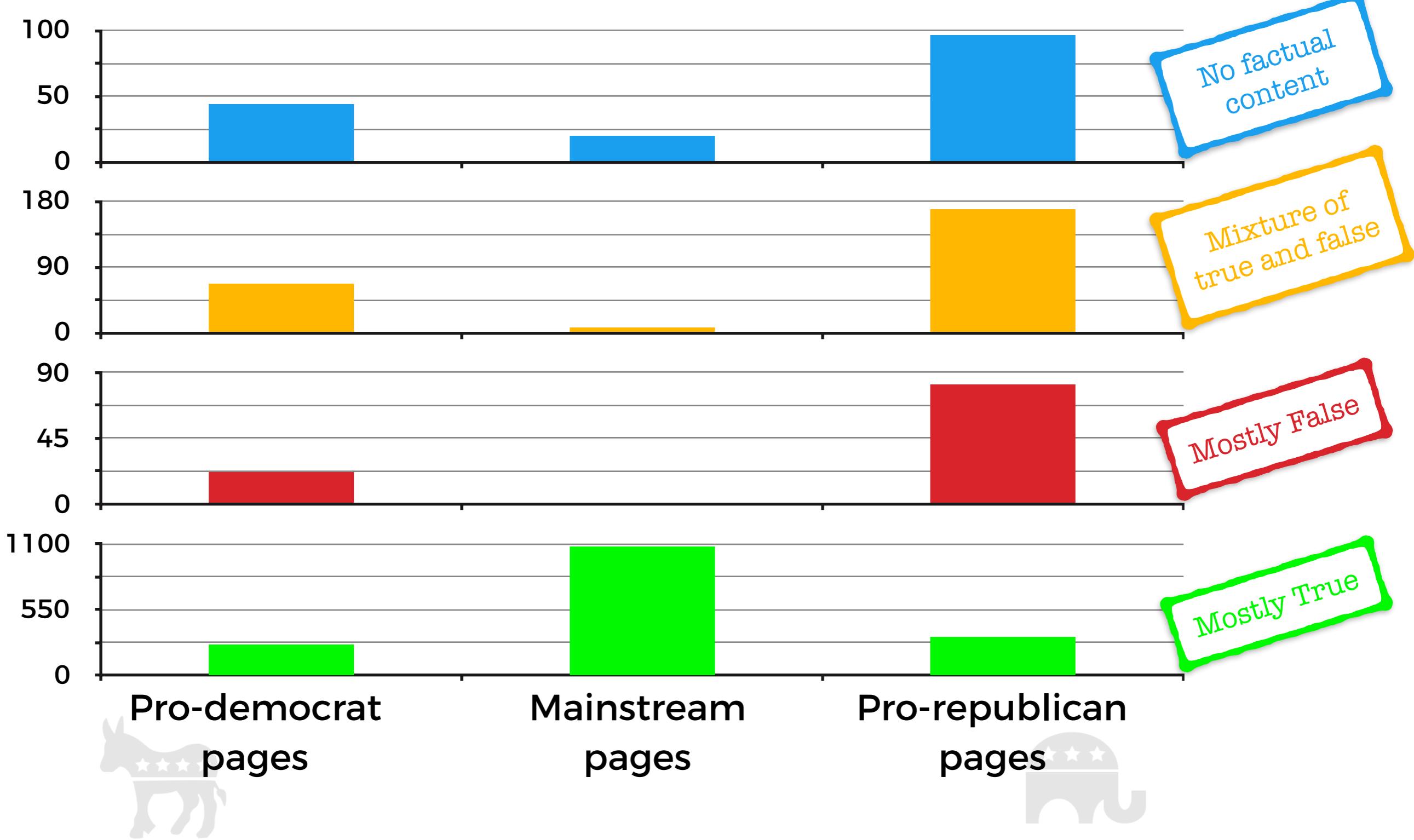
# How much data did we collect?





# How much can we trust these pages?

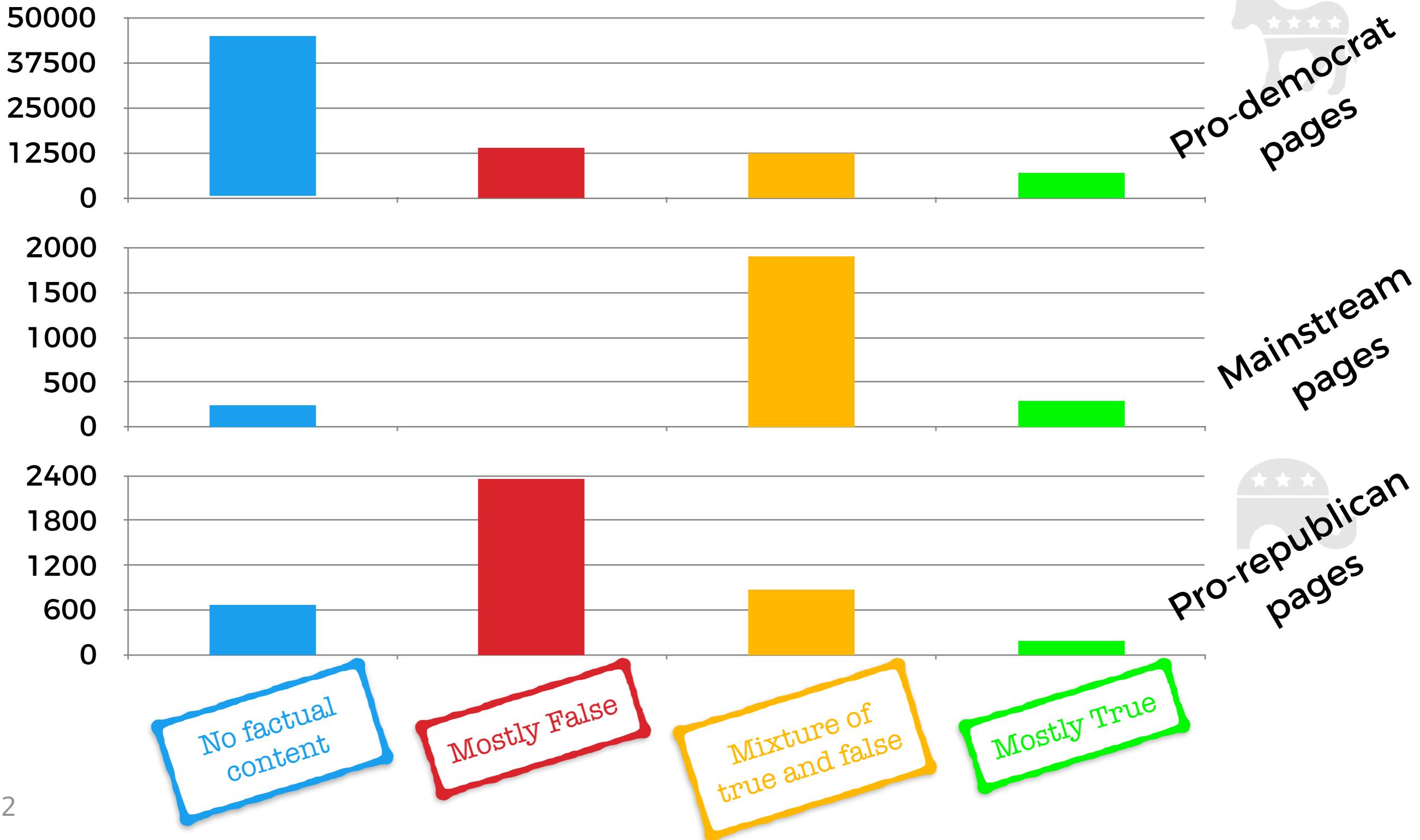
Number of posts published per rating class



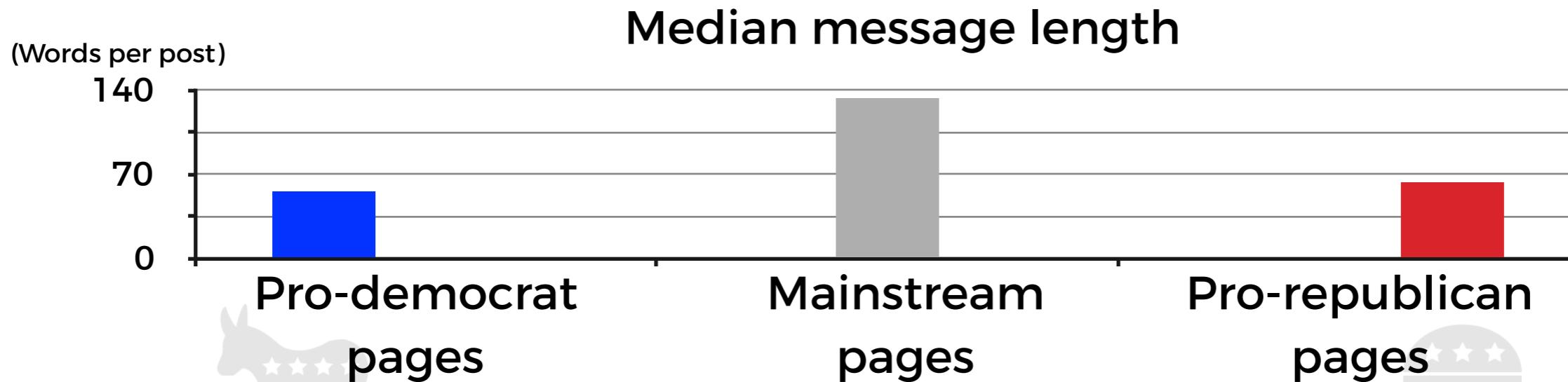


# User engagement metrics

## Median number of shares per post



# What do they say?



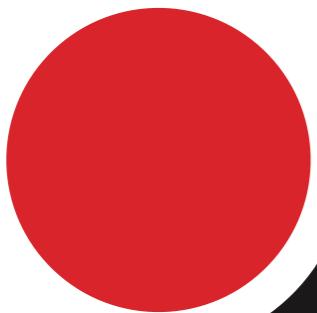
# What do they say?

# Most frequent words in mostly true vs. mostly false messages



# Data Exploitation

Models & Evaluation



# Data & Features

- BuzzFeed dataset

	Messages	Fake	Legitimate
Total	1766	104	1662
Train Set	1236	88	1148
Test Set	530	16	514

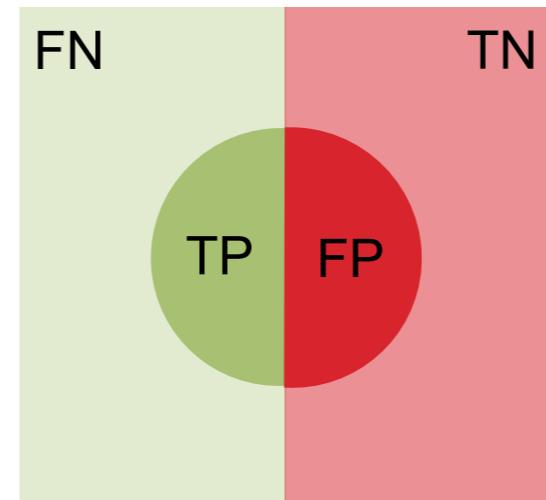
- TF-IDF: 311 features
- Vocabulary Index: 6490 features

# F1-Score

- Accuracy is skewed towards **majority class**
- F1-Score provides a fairer evaluation of **unbalanced** datasets

$$\text{F1-Score} = 2 \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}$$

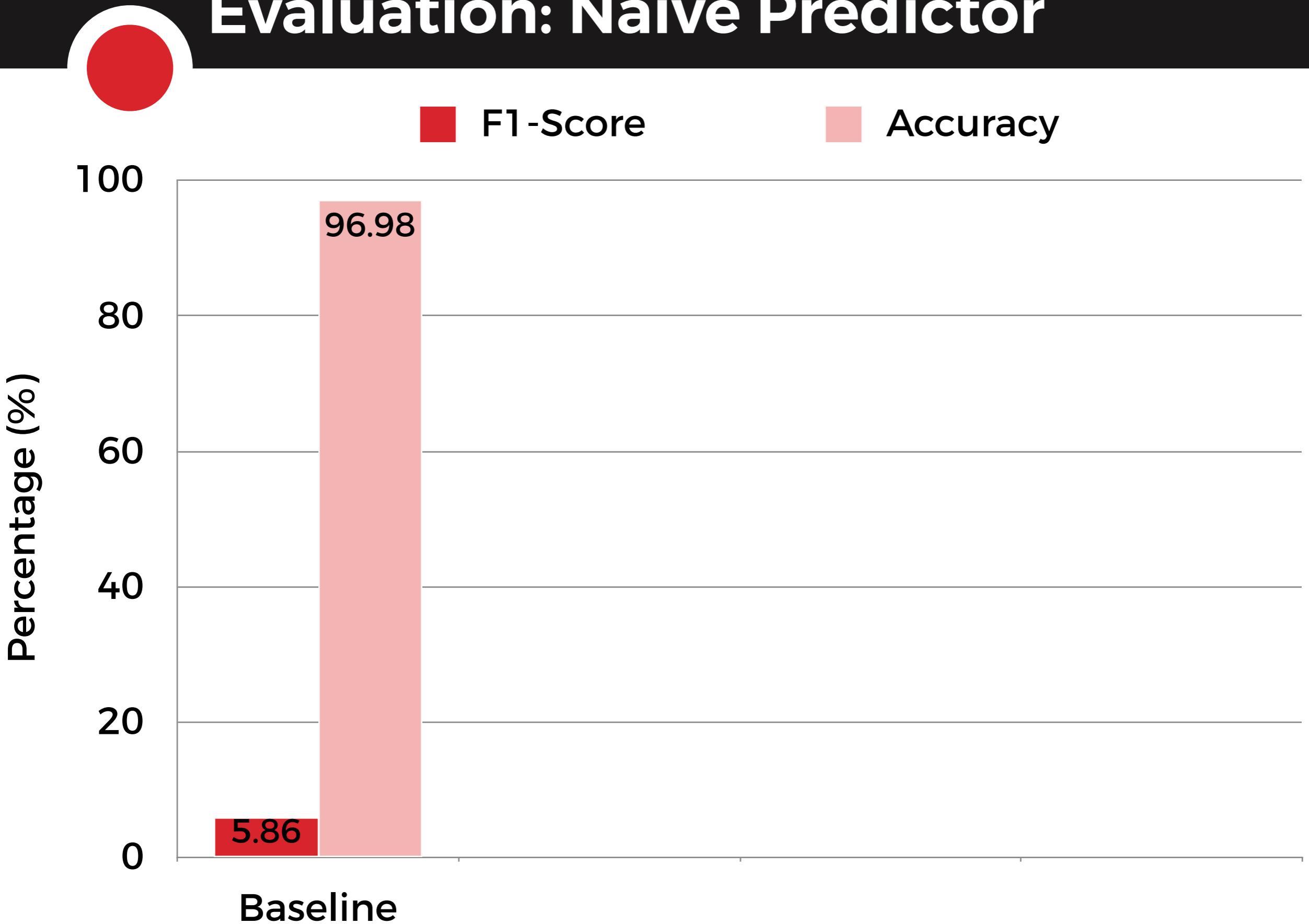
$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$



$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

Positive = Fake  
Negative = Legitimate

# Evaluation: Naive Predictor





# Naive Bayes

- **Features:** Term Frequency - Inverse Document Frequency

$\text{tf-idf}(\text{word}, \text{doc}) = \#(\text{word} \text{ in } \text{doc}) \log [N / \#(\text{doc with word})]$

- Emphasize words **frequent in single document but rare in corpus**

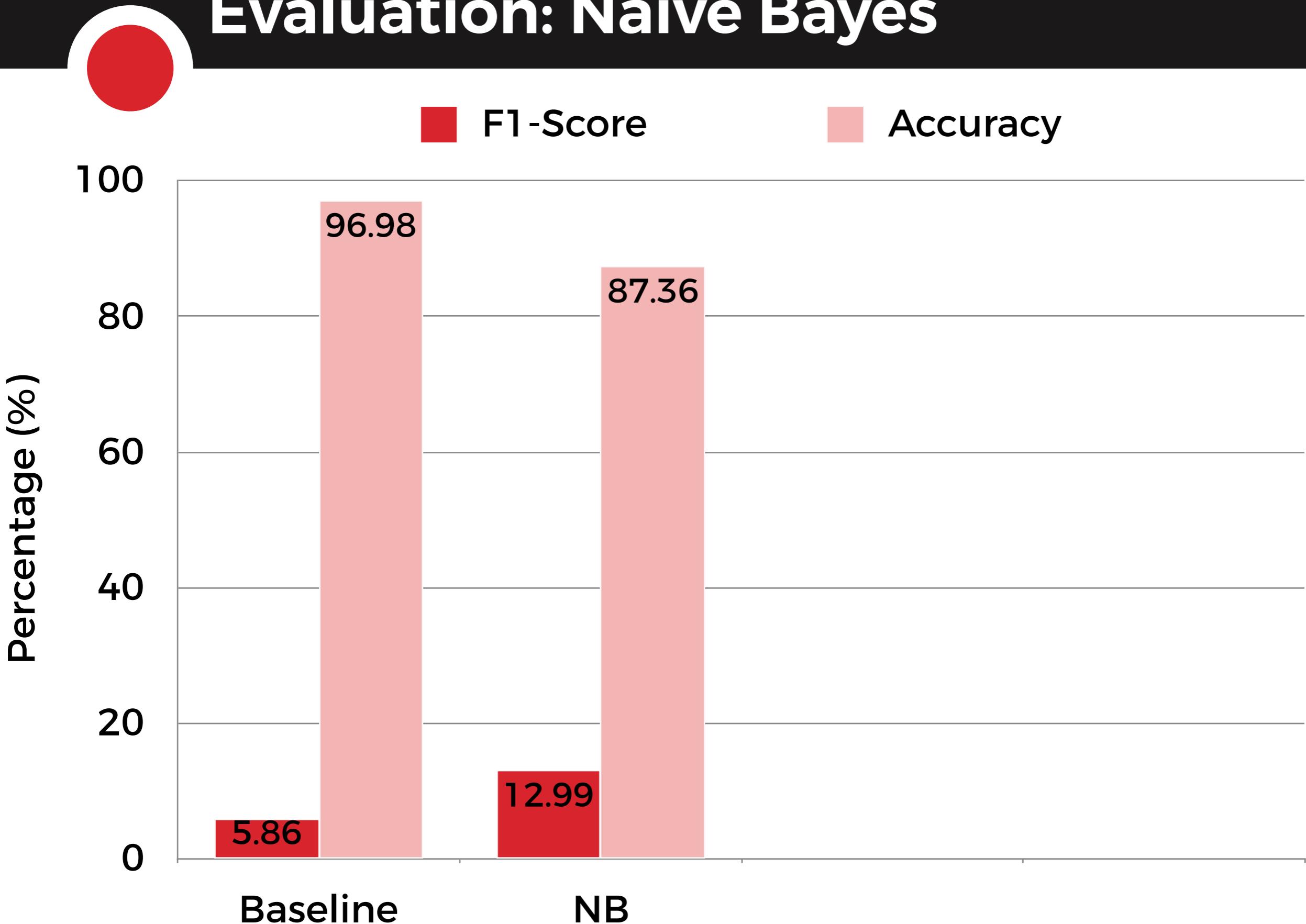
- **Model:** Naive Bayes

- Assumes features are **conditionally independent given class**

- **Widely used** for text classification (before Deep Learning...)

$$P(y|x_1, \dots, x_n) \propto P(y) \prod_{i=1}^n P(x_i|y)$$

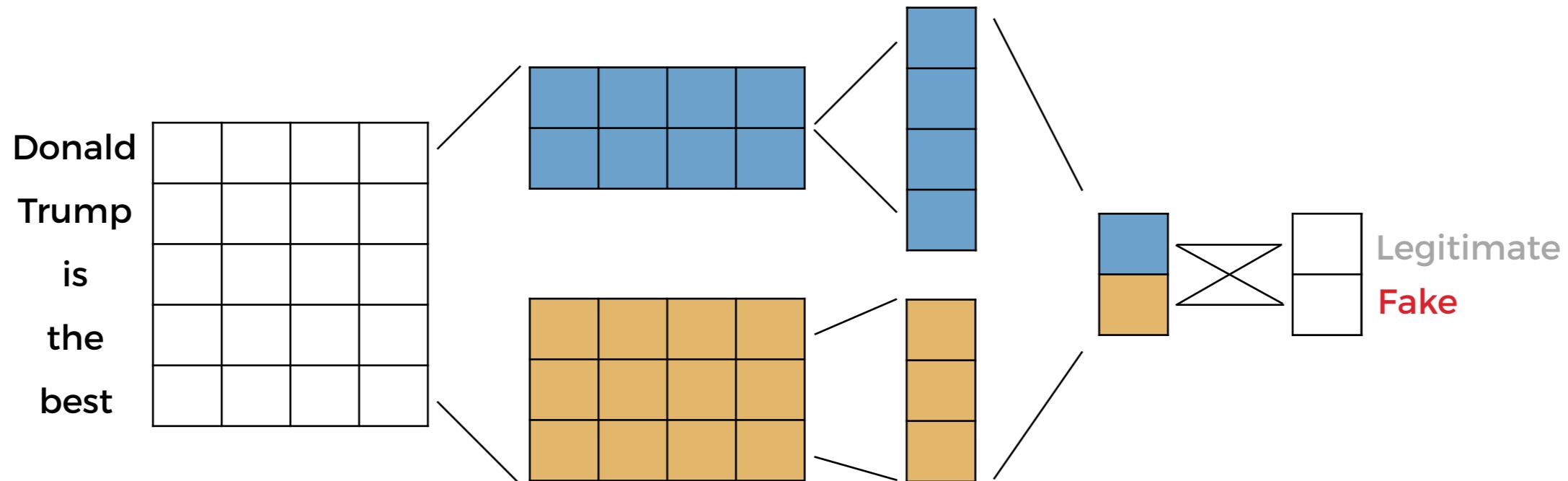
# Evaluation: Naive Bayes



# Convolutional Neural Network

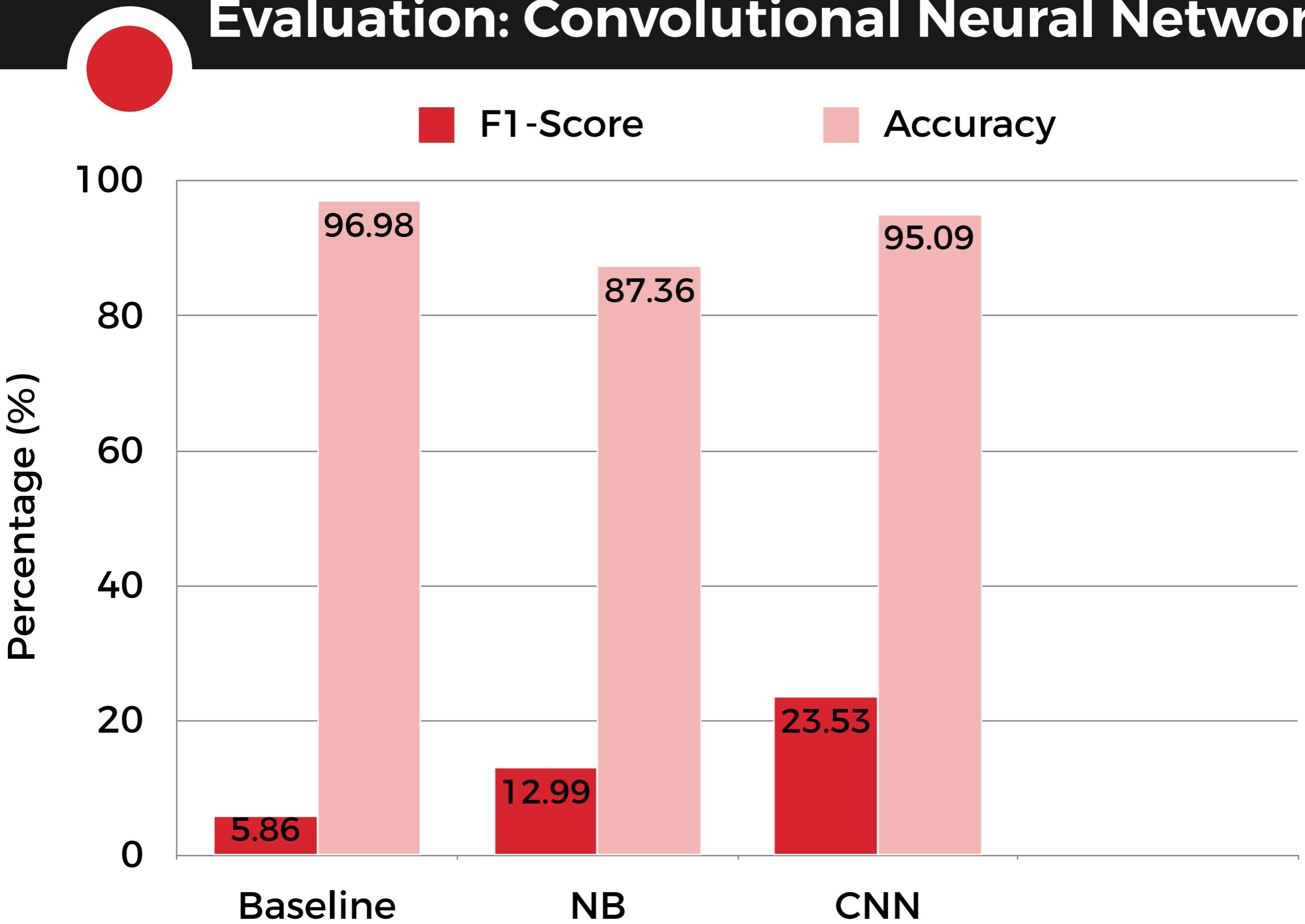
- **Features:** Vocabulary Index
- Each word is represented as a **number**
- Let the neural network **learn the embedding**

## • **Model:** Convolutional Neural Network



- In our case: 2, 3, 4, 5 filter sizes, max-pooling, dropout, L2-regularizer

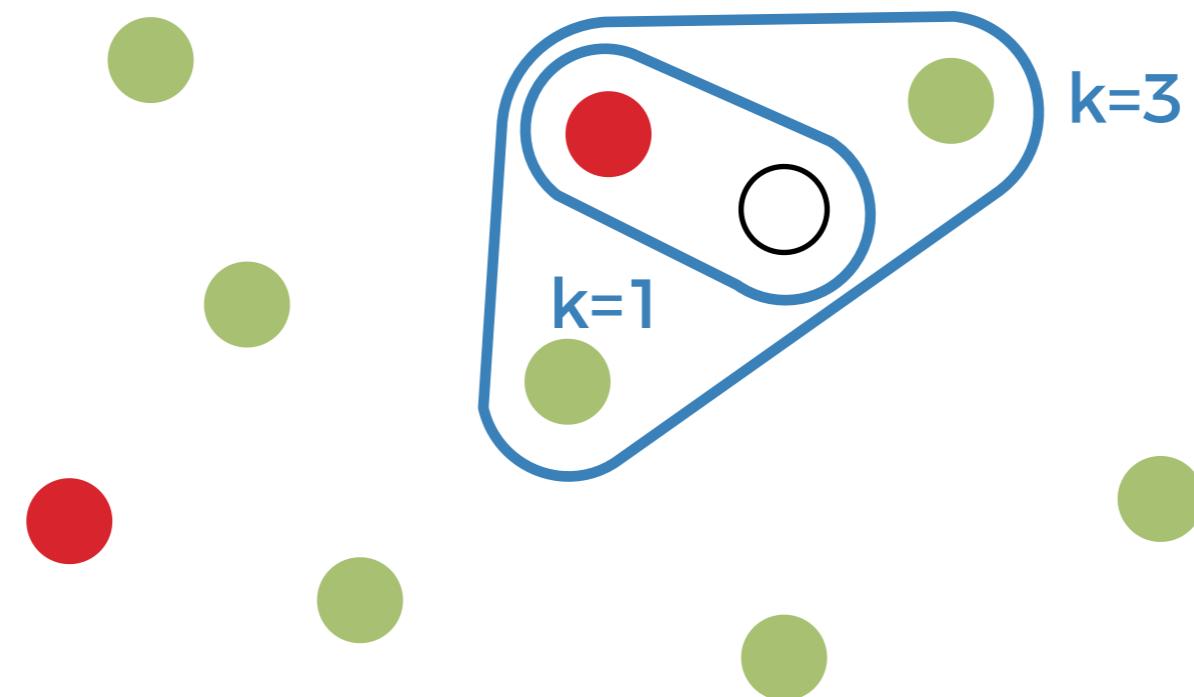
# Evaluation: Convolutional Neural Network



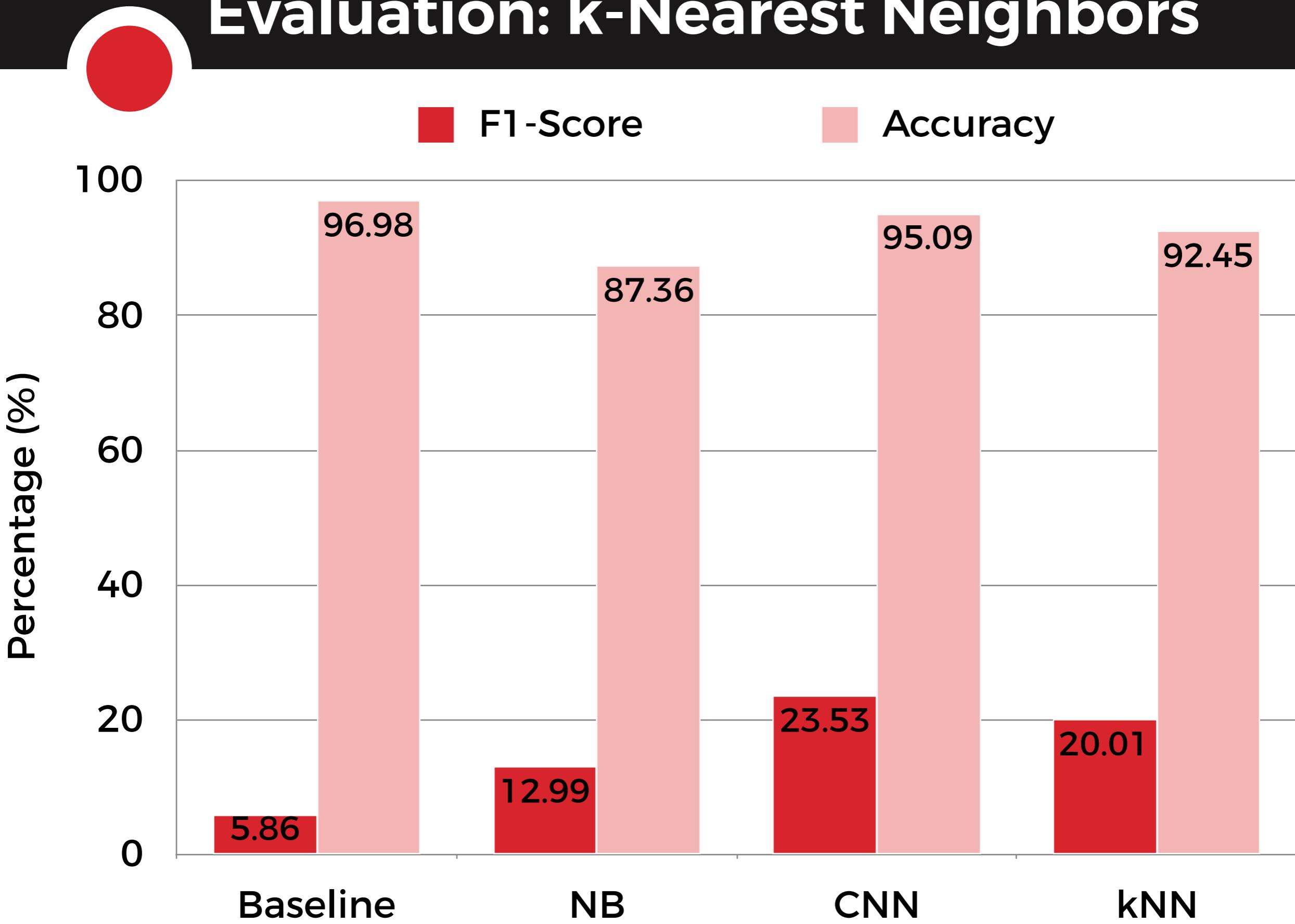


# k-Nearest Neighbours

- **Features:** TF-IDF with non-textual features
- Count of **comments, shares** and **likes**
- **Model:** k-Nearest Neighbours
- **k=1** gives the highest F1-Score
- Probably due to **imbalance** and **number of 'fake' samples**

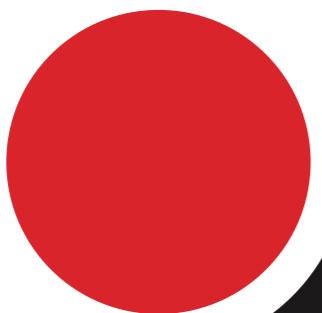


# Evaluation: k-Nearest Neighbors



# **Conclusion**

Further Directions





# Further Directions

- Collect and label more **data**
- Handle better **imbalance** data
- Add more **features**
- Fine-tune **CNN**
- Investigate other **models**

Thank you for your attention!



Do you have (fake/legitimate) questions?

