Hashtag Recommendation System: HOOHLE

ANTONIO COÍN CASTRO
LUIS ANTONIO ORTEGA ANDRÉS
ÁNGEL GABRIEL RIVAS FONFRÍA
IGNACIO SALINAS SÁNCHEZ

Index

01 INTRODUCTION

- Objective and motivation
- Background and state of the art

02 DATA COLLECTION

03 METHODOLOGY

04 EXPERIMENTS

05 CONCLUSIONS AND FUTURE WORK

INTRODUCTION

twitter

@LaraQuijano

Tweets	Following	Followers
74	175	7560

Trends For You

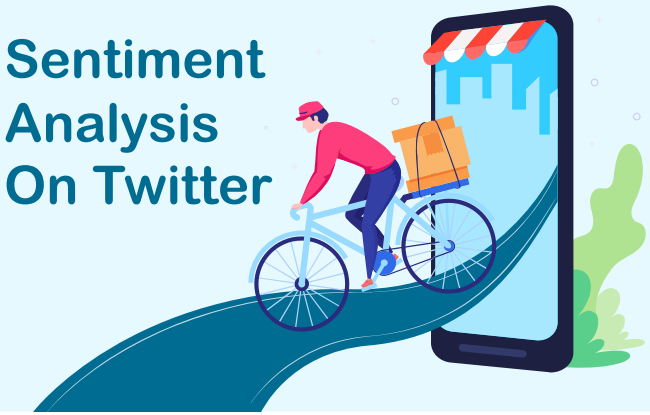
#PICDD
Trending among UAM students.

#Presentations
All the PICDD teachers are tweeting about it.

#Holidays
Coming soon!

What's happening?

Sentiment Analysis On Twitter



PICDD presentations today! Good luck :)

Start Now

For you

Tweets recommended specially for you

Take a look

Who to follow

Refresh

Alvaro Ortigosa
@alvaroOrtigosa

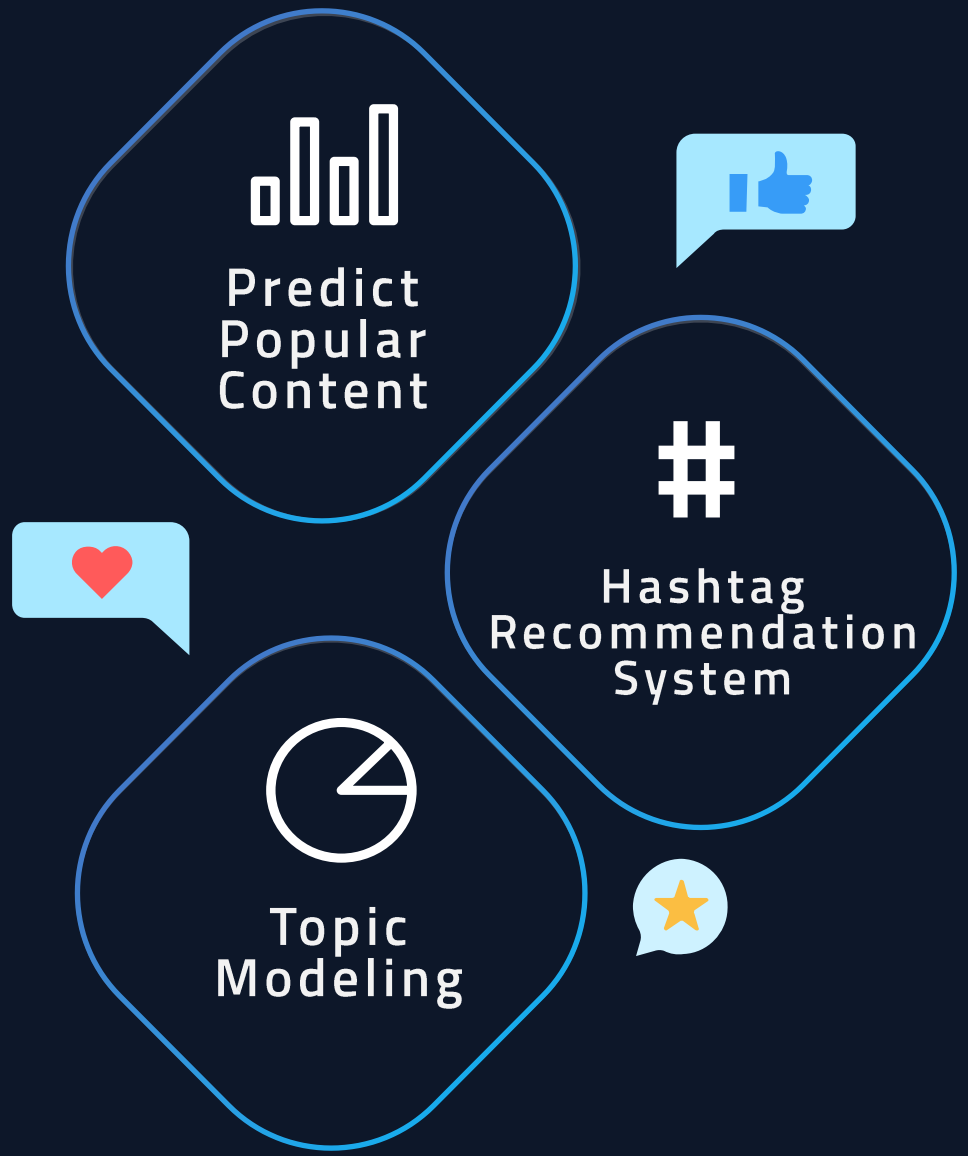
Ruth Cobos
@ruthCobos

Background

Twitter

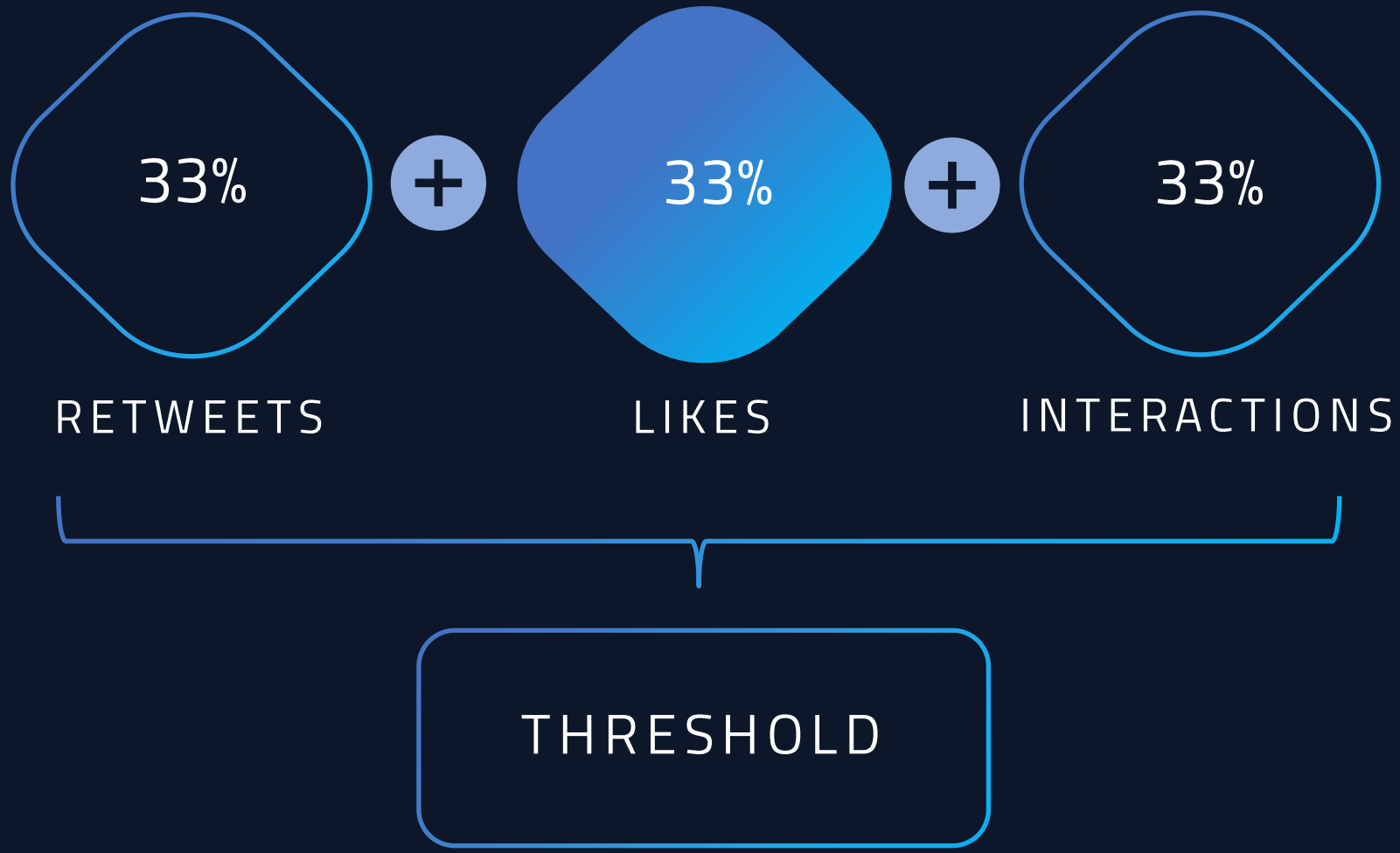
Machine Learning

Modules

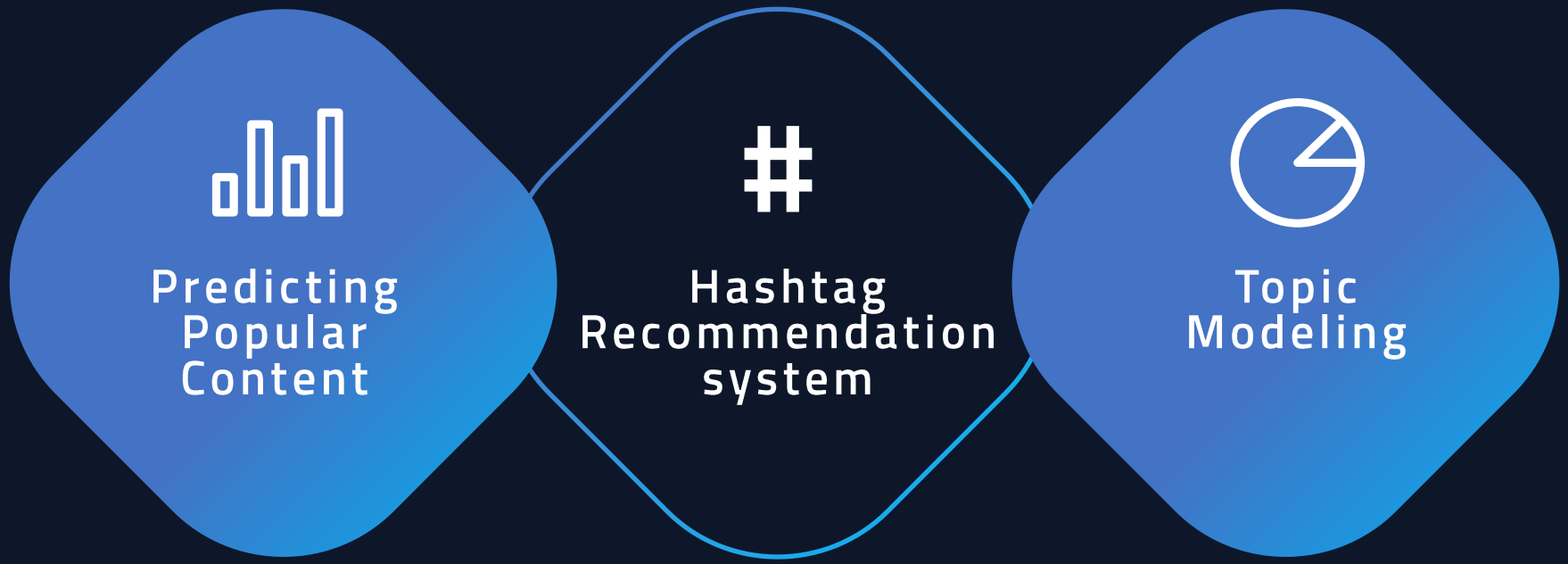

Predict Popular Content
Hashtag Recommendation System
Topic Modeling


Popular #Hashtags Recommender
Write your tweet bellow and we will recommend a Hashtag when you write the symbol #
Qué bien lo pasamos haciendo la asignatura de PICDD #
Universidad
Máster
MUCD
Artículos

Popularity Metric



Current Status



DATA COLLECTION

Existing Datasets

01

CREATED USING THE TWITTER API

02

KAGGLE

03

INDEPENDENT PAGES AND DATASETS

Top Hashtags, Friend or Follow or Twitter SEISMIC.

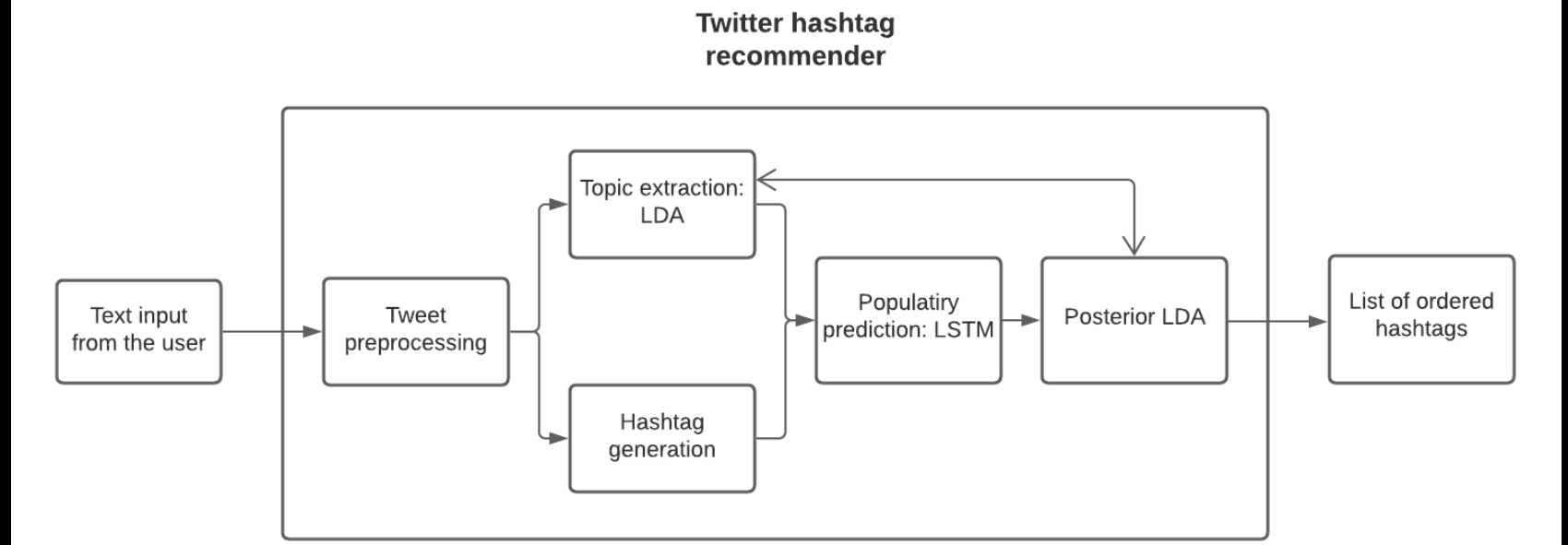
Option Chosen

1. Initial Training Period
   - Kaggle
   - Twitter SEISMIC
   - Internet Archive
2. Select the threshold
3. **Maintenance + improvement phase**

METHOLODOGY

Hoohle Flowchart

Tweet Preprocessing + Topic Extraction + Existent Hashtag Retrieval
+ New Hashtag Generation + Popularity Prediction



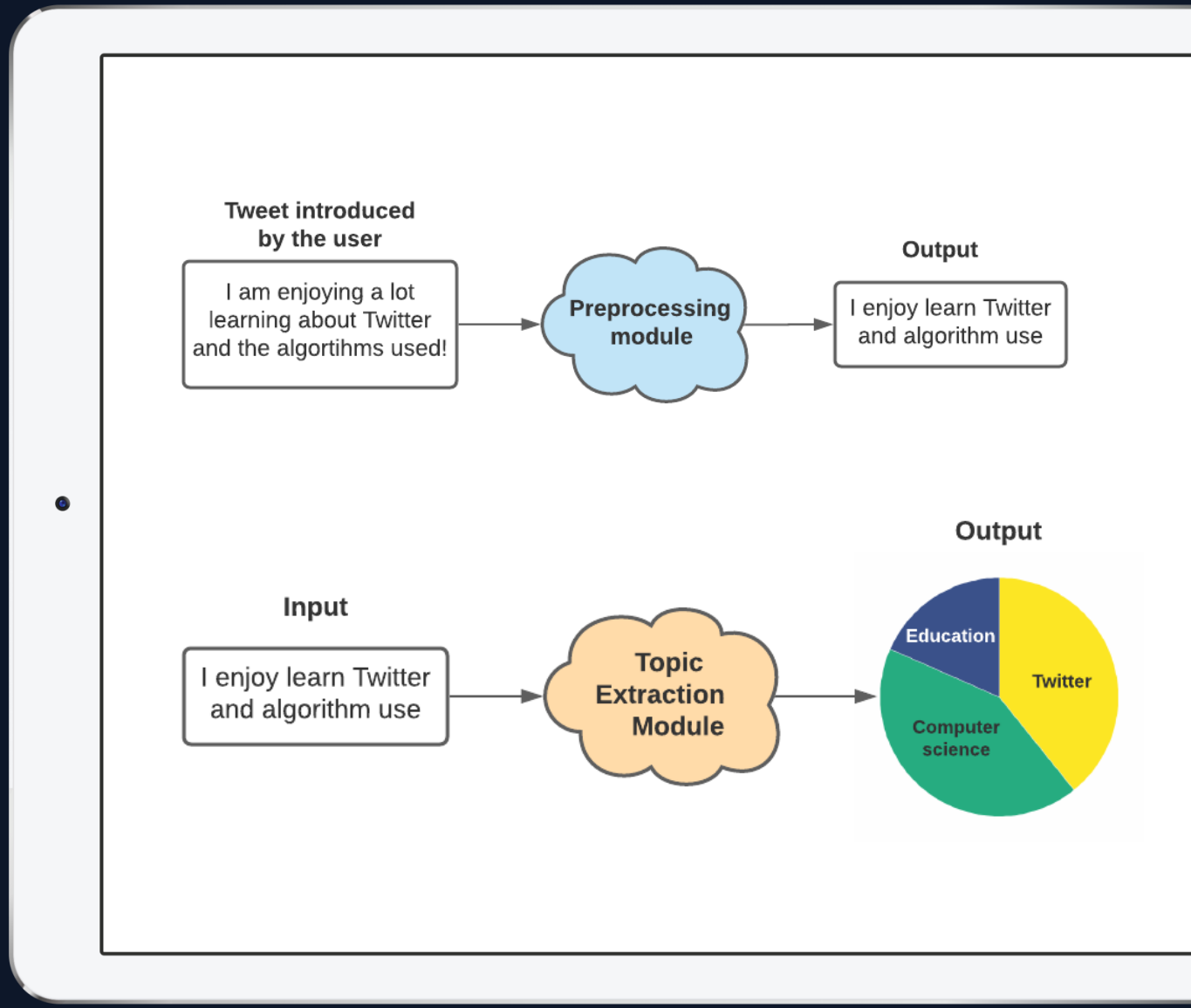
HOOHLE Flowchart

STEP 1

Tweet Preprocessing

STEP 2

Topic Extraction Module

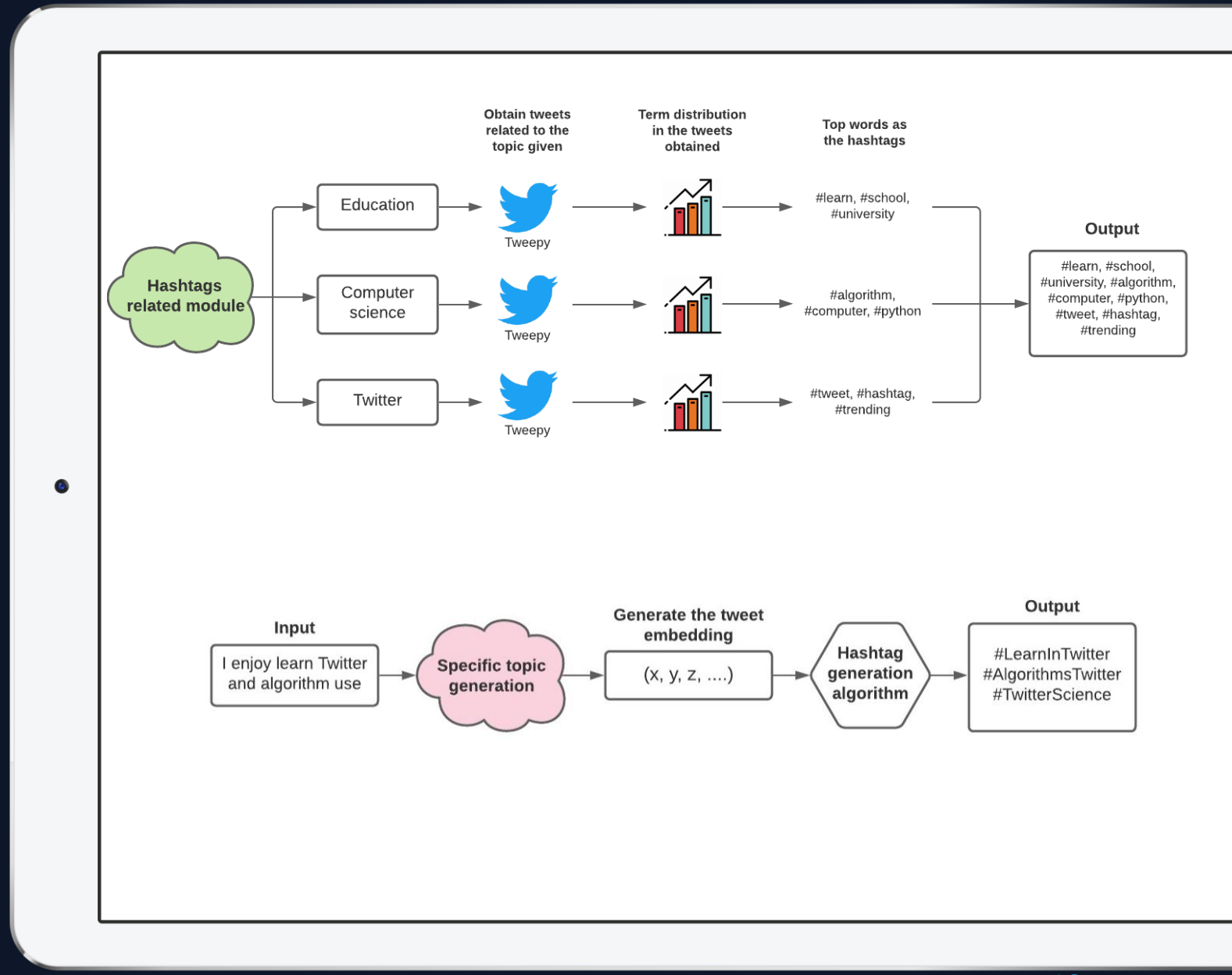


STEP 3-1

Existent hashtags retrieval

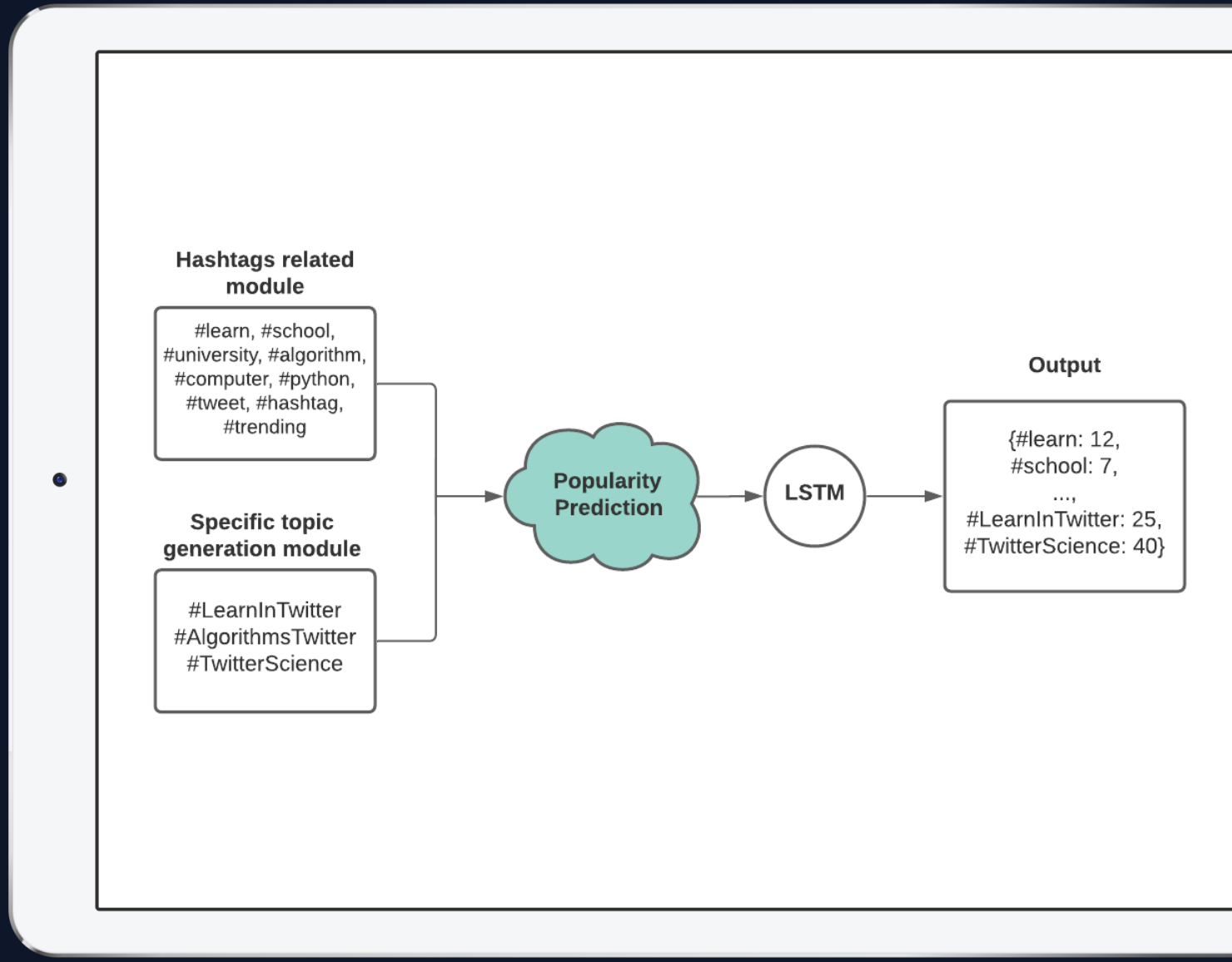
STEP 3-2

New hashtag generation



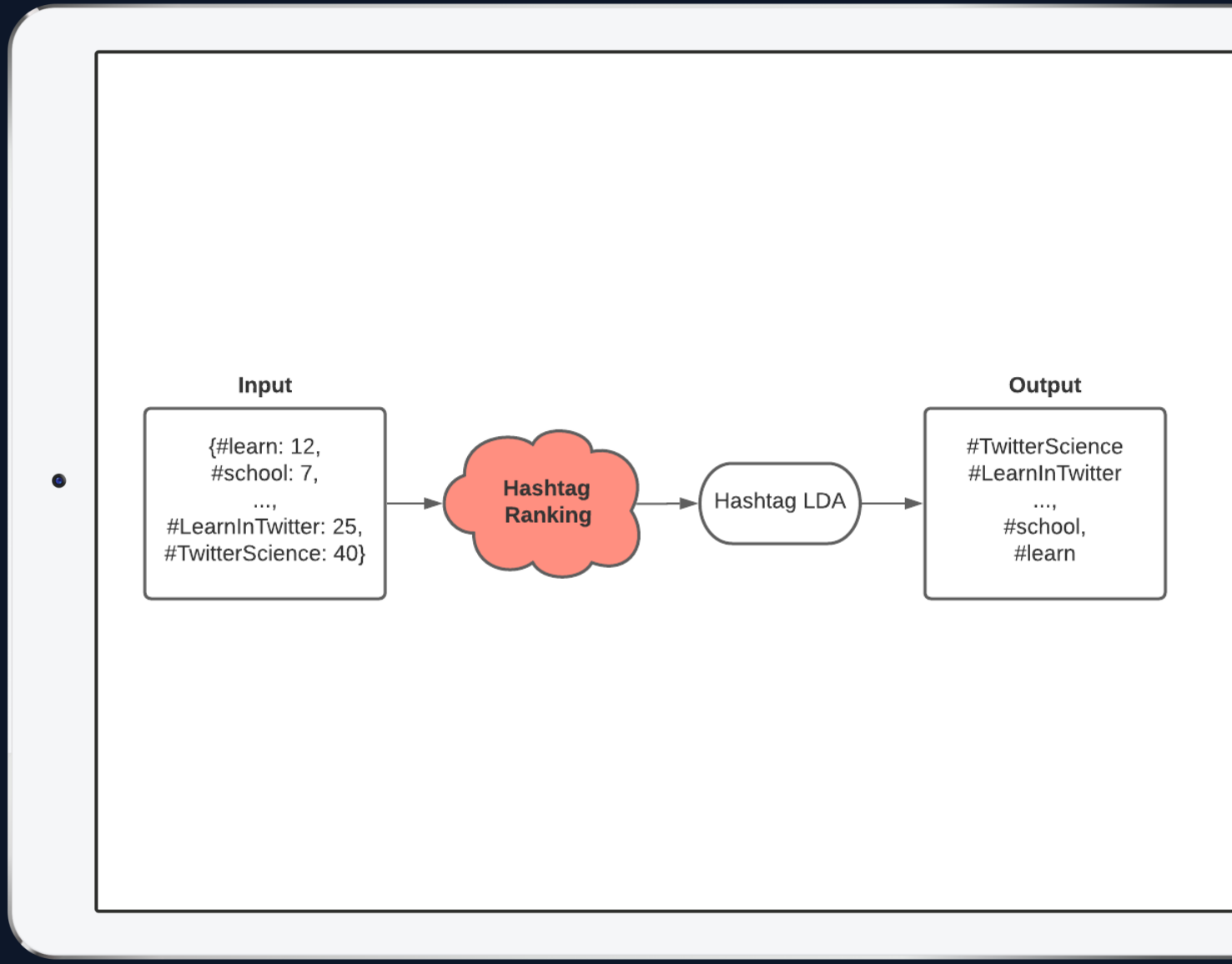
STEP 4

Popularity Prediction

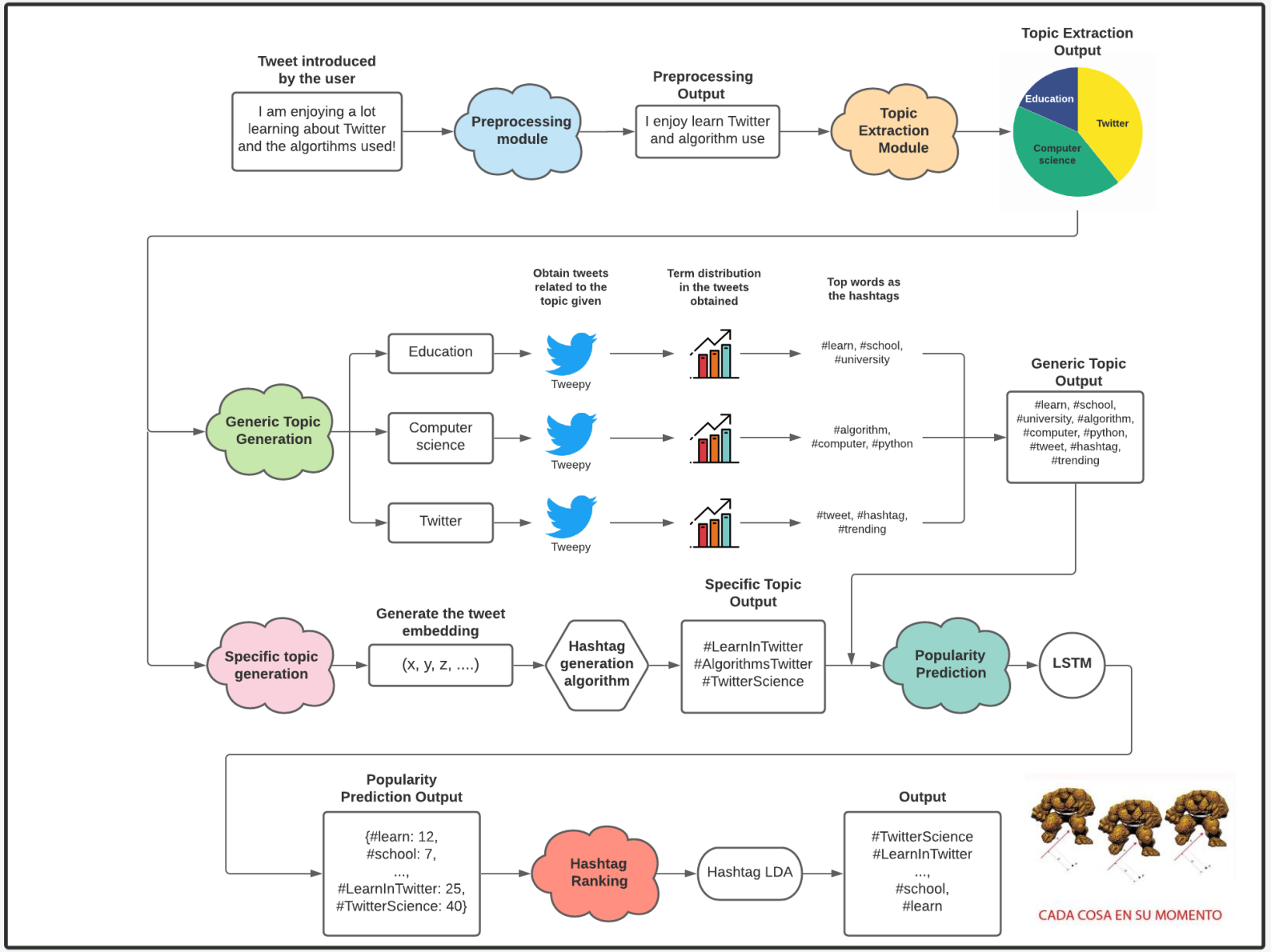


STEP 5

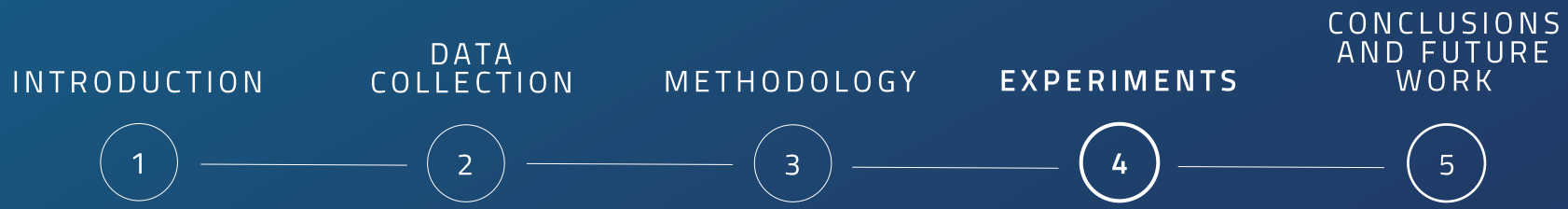
Proposal Validation



Workflow


Tweet introduced by the user
I am enjoying a lot learning about Twitter and the algortihms used!
Preprocessing module
Preprocessing Output
I enjoy learn Twitter and algorithm use
Topic Extraction Module
Topic Extraction Output
Education
Twitter
Computer science
Obtain tweets related to the topic given
Term distribution in the tweets obtained
Top words as the hashtags
Generic Topic Generation
Education
Computer science
Twitter
Tweepy
Tweepy
Tweepy
#learn, #school, #university
#algorithm, #computer, #python
#tweet, #hashtag, #trending
Generic Topic Output
#learn, #school, #university, #algorithm, #computer, #python, #tweet, #hashtag, #trending
Specific topic generation
Generate the tweet embedding
(x, y, z, ...)
Hashtag generation algorithm
Specific Topic Output
#LearnInTwitter
#AlgorithmsTwitter
#TwitterScience
Popularity Prediction
LSTM
Popularity Prediction Output
{#learn: 12,
#school: 7,
...,
#LearnInTwitter: 25,
#TwitterScience: 40}
Hashtag Ranking
Hashtag LDA
Output
#TwitterScience
#LearnInTwitter
...,
#school,
#learn
CADA COSA EN SU MOMENTO

EXPERIMENTS



Validation of the system

SUPERVISED METRIC

F1-score for popularity prediction

SIMILARITY METRIC

Popularity-relevance trade-off

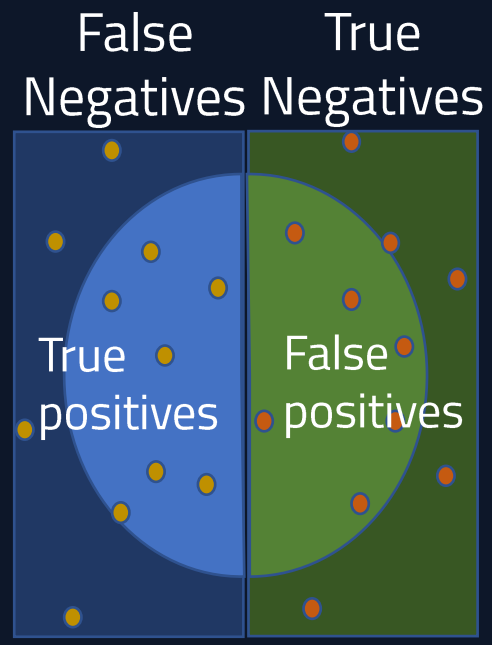
TEMPORAL METRIC

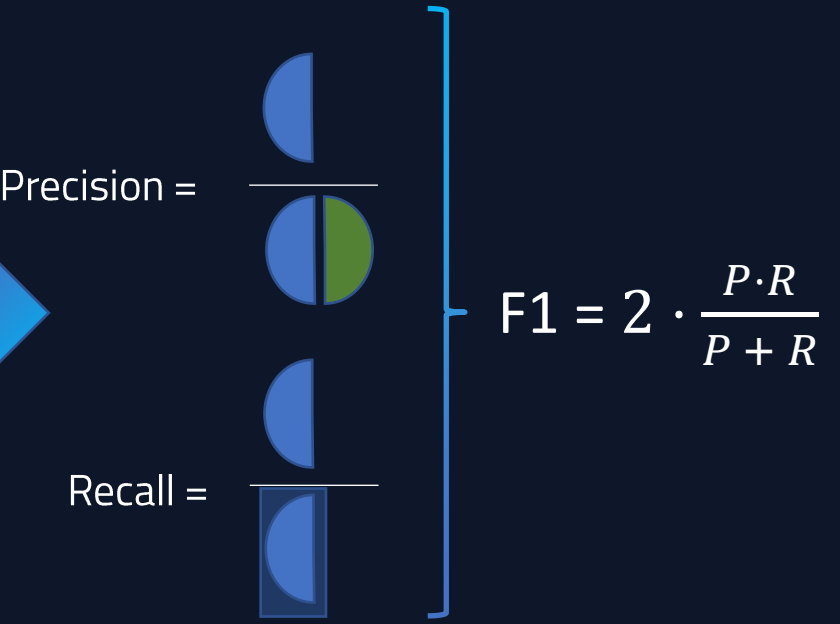
Measure impact over time

Validation of the system

SUPERVISED METRIC

F1-score for popularity prediction





Validation of the system

 # SIMILARITY METRIC

Popularity-relevance trade-off

$\tilde{p} = LDA(\#+$  $)$

$p = LDA($  $)$

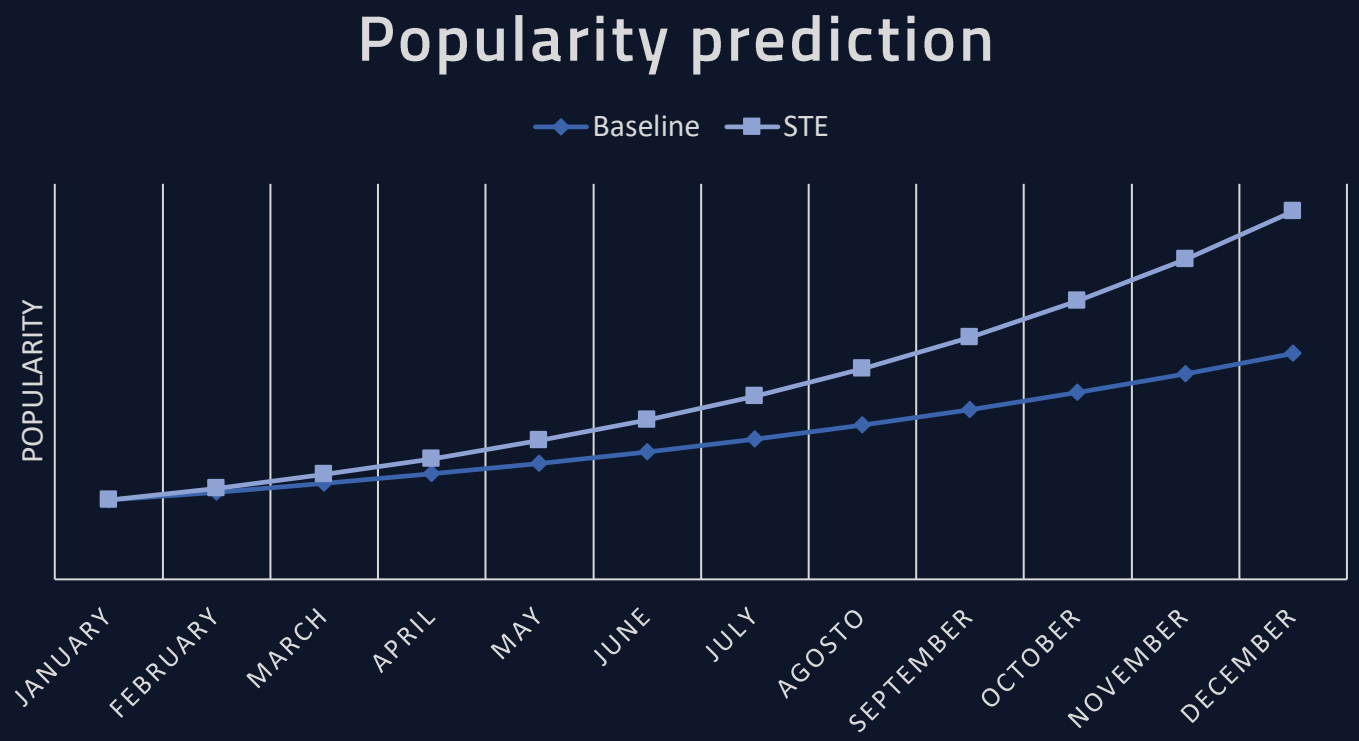
$$l(p, \tilde{p}) = \sum_{i=0}^{topics} |p_i - \tilde{p}_i|$$

THRESHOLD

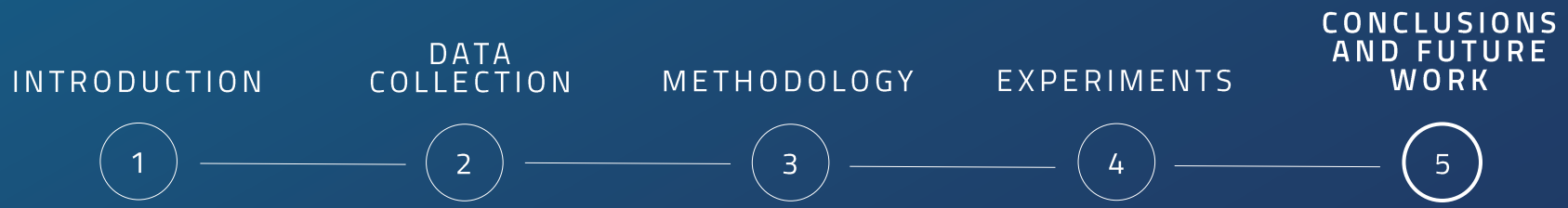
Validation of the system

TEMPORAL METRIC

Measure impact over time



CONCLUSIONS AND FUTURE WORK



Conclusions

End to end

Hashtag recommendation system

Previous Topic Related Hashtags + Hashtag Generation

Popularity Prediction System

33% + 33% + 33%

RETWEETS LIKES INTERACTIONS

THRESHOLD

Future Work

01 Analyze hyperlinks and media attached to the tweet.

02 Adjust the model to other social networks.

03 Use contextual data such as mentions.


"That's all Folks!"

Questions?