



MOVIE RECOMMENDER SYSTEM

SWAPNIL SAH
ADITYA SHARMA
ABHISHEK BASU

19BCE7458
19BCD7165
19BCE7326

GUIDED BY: PROF.HARI SEETHA

Contents

1

Originality

How we made this project our own and different from the rest

4

Results

Our final outcomes

2

Methodology

How we approached the idea and worked towards the final result

5

Objectives met

3

Data Cleaning and Pre-processing

An inner look at our data and its workings

6

Applications

where can be used



Originality

Source: TMDB and Kaggle

Before pre-processing -

Number of rows:4806

Number of cols:20

After pre-processing -

Number of rows:4806

Number fo cols:3

After adding Bollywood data -

Number fo cols:8296

Number fo cols:3

Pre-Processing

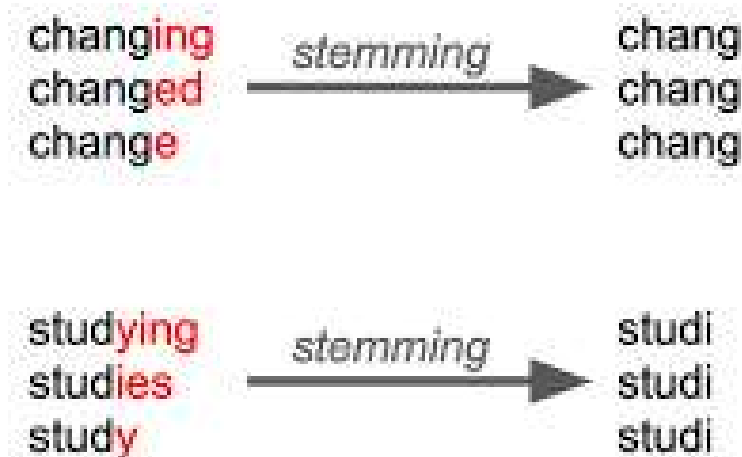
BEFORE

Dropped irrelevant features

We have dropped the features apart from genre, overview, title, cast and crew and combined them into one feature named 'tags'

Used stemming

Used for pre processing of textual data by trimming the last few characters of similar words



movies.head(2)

	budget	genres	homepage	id	keywords	original_language	original_title	overview	popularity	pro
0	237000000	[{"id": 28, "name": "Action"}, {"id": 12, "name": "Adventure"}]	http://www.avatarmovie.com/	19995	[{"id": 1463, "name": "culture clash"}, {"id": 1464, "name": "culture clash"}]	en	Avatar	In the 22nd century, a paraplegic Marine is dispatched to the moon Pandora on a unique mission, but becomes torn between following orders and protecting those who have become his family.	150.437577	
1	300000000	[{"id": 12, "name": "Adventure"}, {"id": 14, "name": "Fantasy"}]	http://disney.go.com/disneypictures/pirates/	285	[{"id": 270, "name": "ocean"}, {"id": 726, "name": "ocean"}]	en	Pirates of the Caribbean: At World's End	Captain Barbossa, long believed to be dead, has returned to the Caribbean Sea.	139.082615	

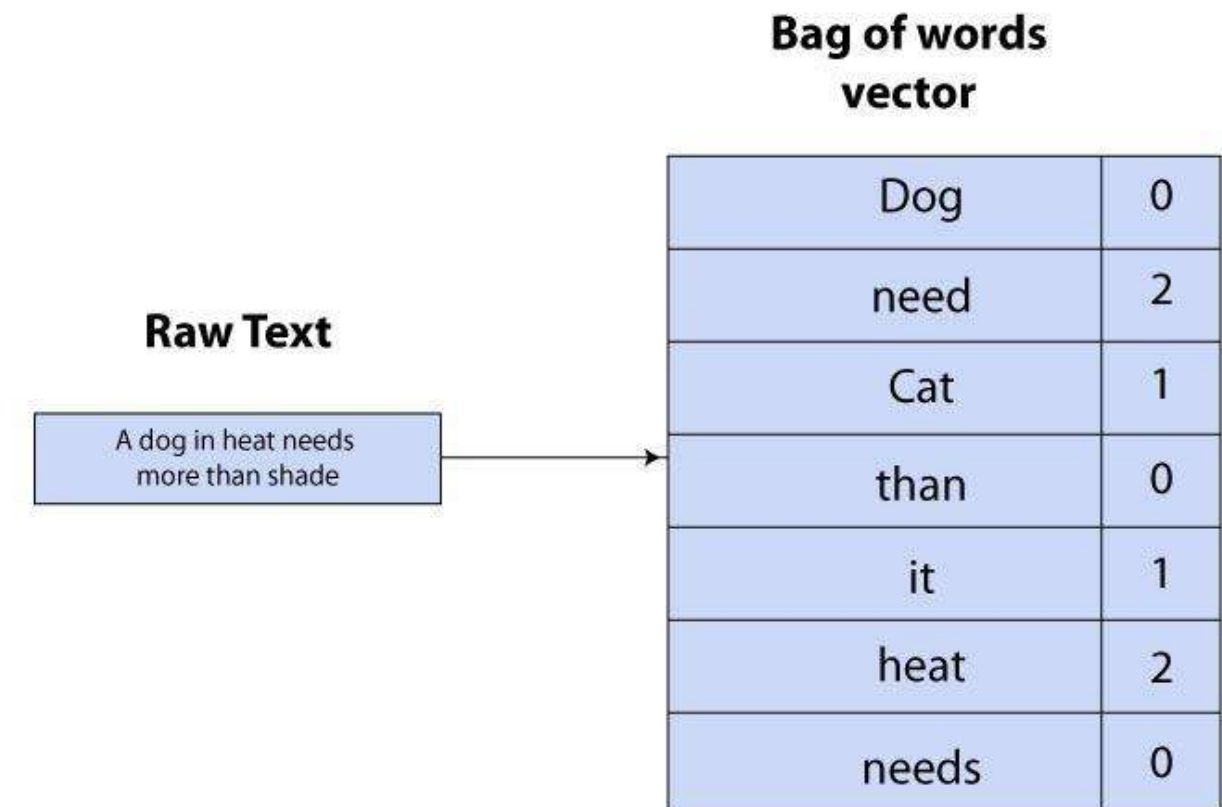
AFTER

[35] new.head()

	movie_id	title	tags
0	19995	Avatar	in the 22nd century, a paraplegic marine is dispatched to the moon pandora on a unique mission, but becomes torn between following orders and protecting those who have become his family.
1	285	Pirates of the Caribbean: At World's End	captain barbossa, long believed to be dead, has returned to the caribbean sea.
2	206647	Spectre	a cryptic message from bond's past sends him on a hunt for a secret organization that could threaten the world.
3	49026	The Dark Knight Rises	follow the death of district attorney harvey dent and the rise of a new criminal force in batman's Gotham City.
4	49529	John Carter	john carter is a war-weary, former military hero, afflicted with an incurable degenerative disease.

Methodology

- We have used **Content-based filtering**
- This compares and recommends other movies similar to the selected movie
- $f(\text{movie}) \rightarrow \{\text{movies}\}$
- We have done preprocessing and stemming
- We have used **CountVectorizer** to convert textual data to numerical data
- This assisted in calculating the distances between different titles
- We use **Cosine similarity** as a similarity measure to quantify how similar the movies are based on their similarities of different properties.
- Mathematically, it shows the cosine of the angle of two vectors projected in a multidimensional space which in turn helps to find similar objects.



Model Created

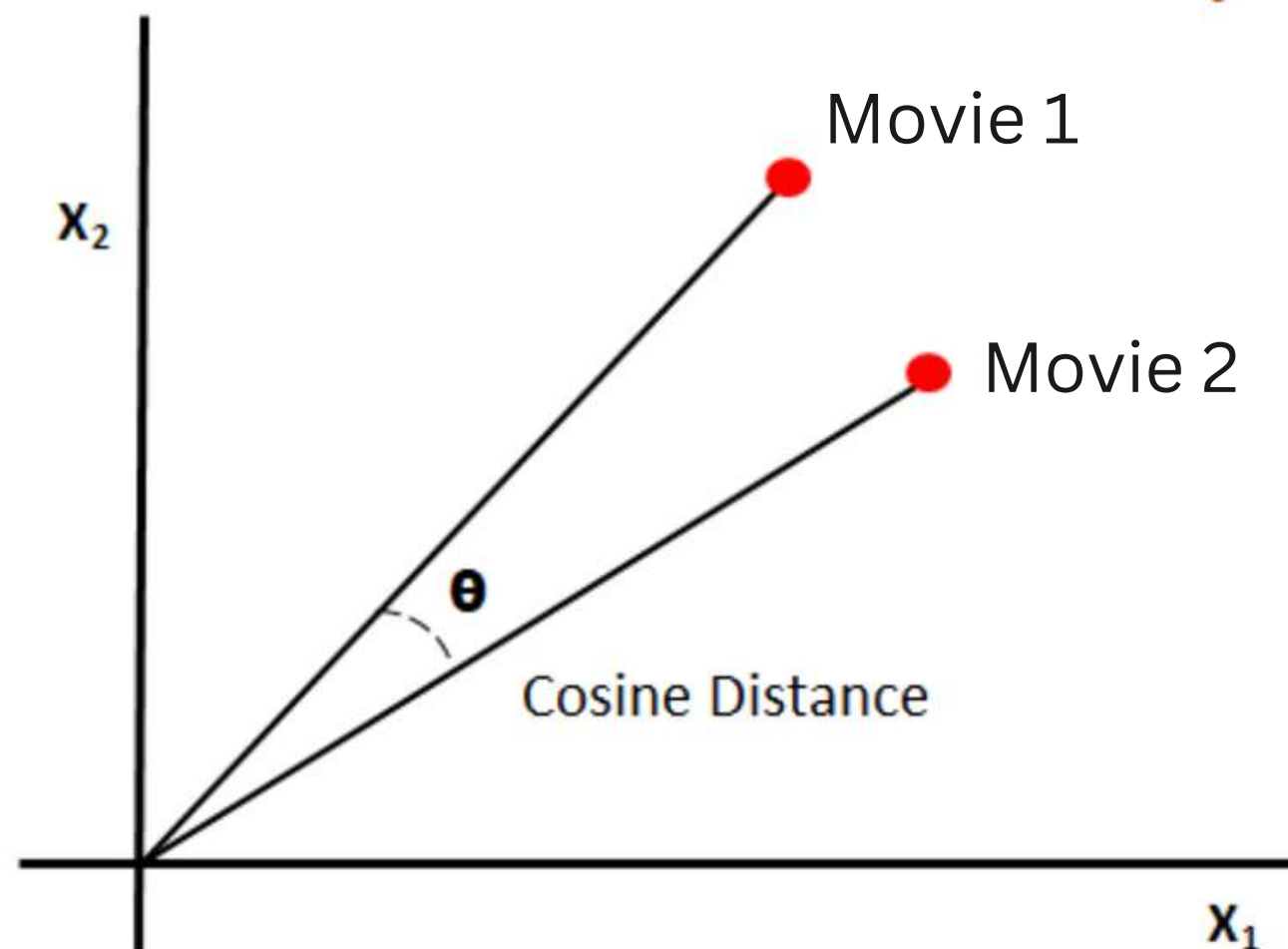
Count Vectorizer

Breaking down a sentence or any text into words by performing preprocessing tasks like converting all words to lowercase, thus removing special characters

Cosine Similarity

Cosine similarity is a measure of similarity between two non-zero vectors of an inner product space that measures the cosine of the angle between them. nltk toolkit module is used in this program.

Cosine Distance/Similarity



Test Cases

✓
0s



```
recommend('Batman')
```

Batman
Batman & Robin
Batman Begins
Batman Returns
The R.M.

✓
1s

```
[55] recommend('Avatar')
```

Aliens vs Predator: Requiem
Aliens
Falcon Rising
Independence Day
Titan A.E.

✓
1s



```
recommend("Pirates of the Caribbean: At World's End")
```

Pirates of the Caribbean: Dead Man's Chest
Pirates of the Caribbean: The Curse of the Black Pearl
Pirates of the Caribbean: On Stranger Tides
Life of Pi
20,000 Leagues Under the Sea

✓
0s



```
recommend('The Avengers')
```

Iron Man 3
Avengers: Age of Ultron
Captain America: Civil War
Captain America: The First Avenger
Iron Man

RESULTS

PRE-PROCESSING

We were able to transform the data into a format that is more easily and effectively processed for our model to be applied without difficulty.

RECOMMENDATION

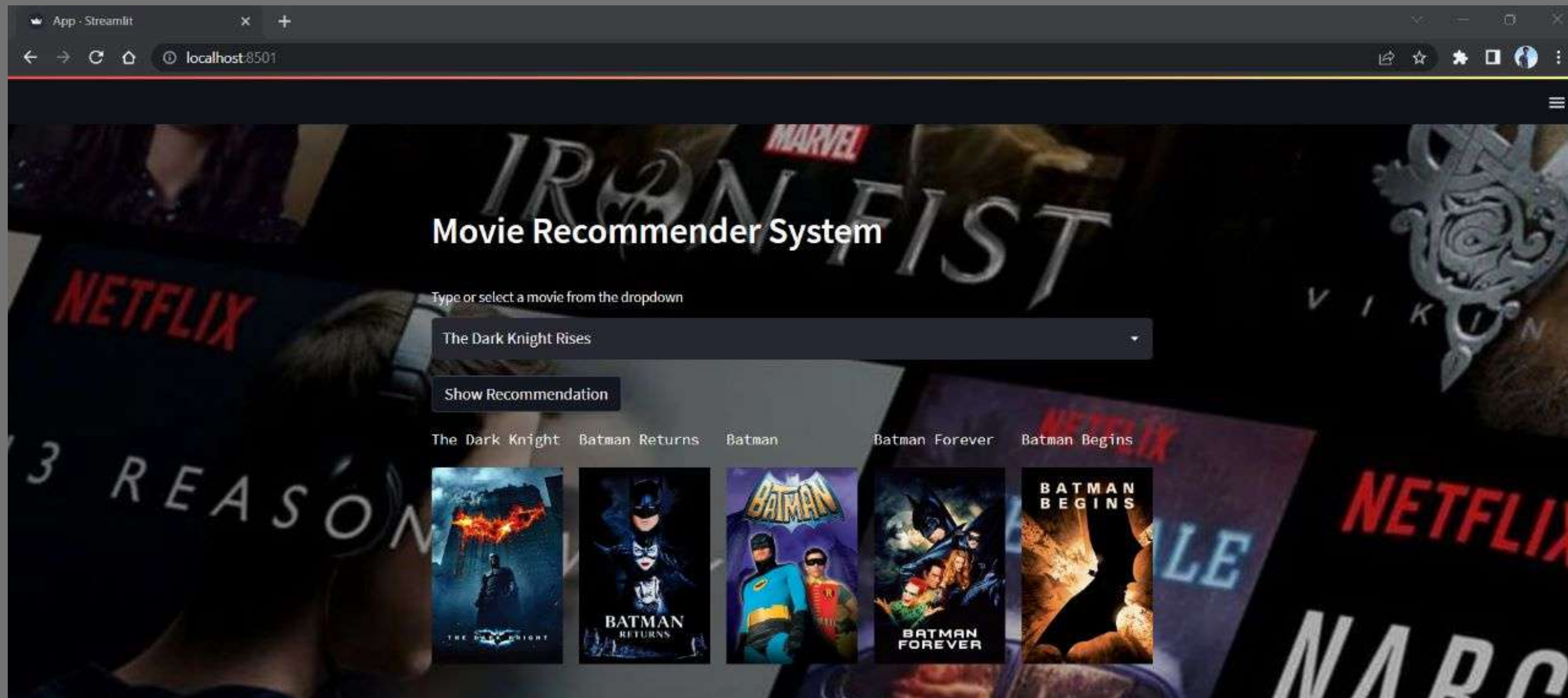
We successfully recommended movies based on their respective titles.
e.g. Batman gave the recommendation of - Batman & Robin, Batman Begins, and Batman Returns

MODEL USED

Model used for the execution of our program was content based filtering which was selected to ensure that the results are closely related to the user input to give more accurate results.

OBJECTIVE MET

Our Movie Recommender System is able to predict movies of similar types. It predicted only those movies that the user will most likely watch.



APPLICATIONS

The interface created will be helpful in the process of assisting users in their recommendation with the help of the movie they enter.

This could be beneficial for e-commerce, retail, media, banking and telecom among others.

This leads to increased sales/conversion, user satisfaction and reduced churn.
Companies that already use such systems are Amazon, Netflix, Spotify, LinkedIn etc

TIMELINE

Milestones:

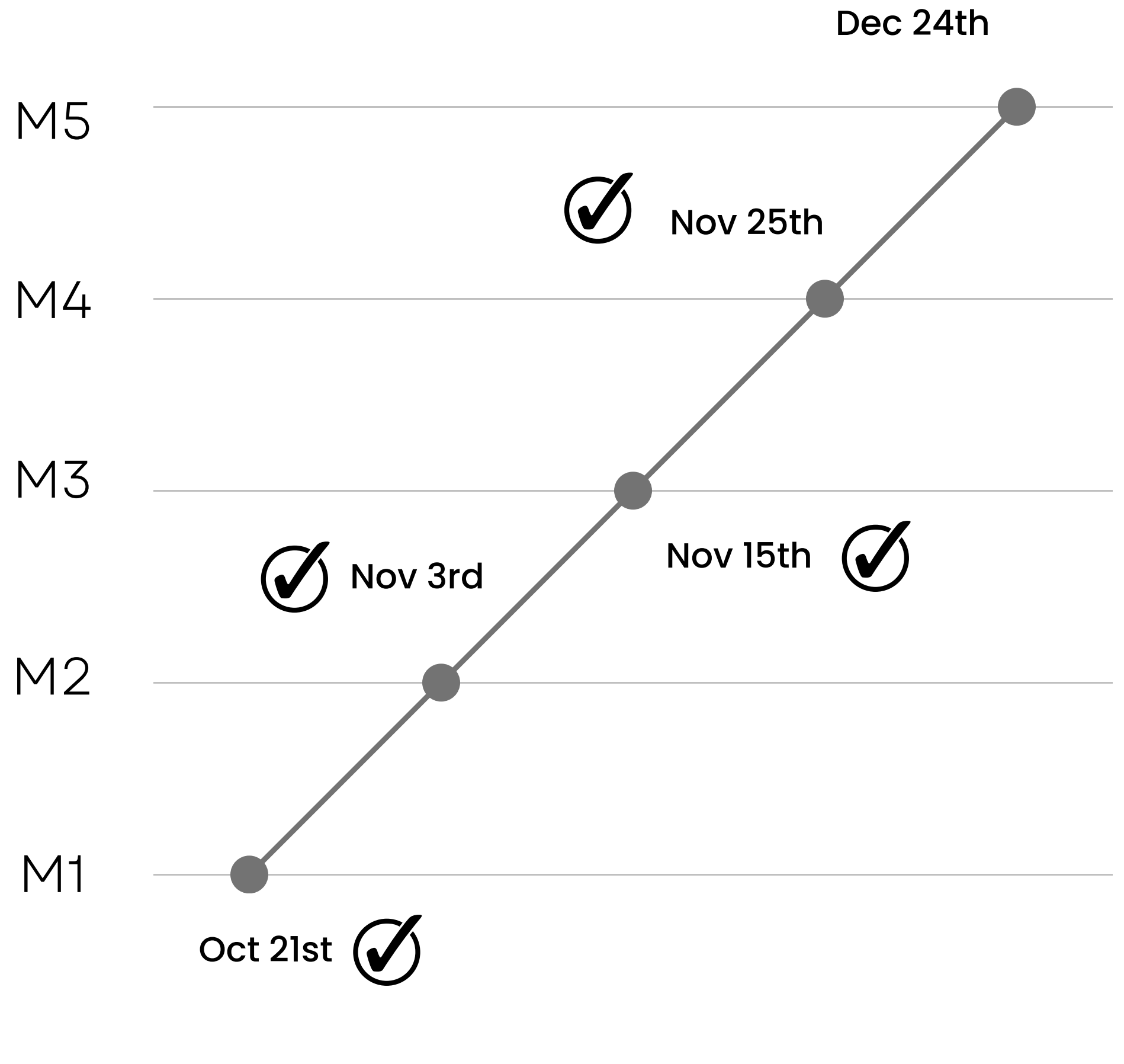
M1. Dataset identification

M2. Data cleaning / Data pre-processing

M3. Model development

M4. Designing and implementation

M5. Final submission





THANK YOU