

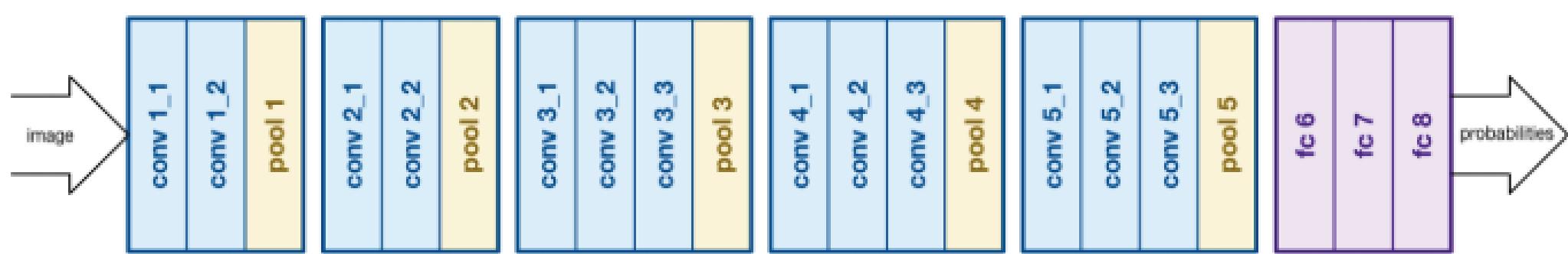
# StoryReel : Generating Short Stories From Celebrity Images

## ABSTRACT

Language or story generation has a long history of pursuit in artificial intelligence. We present a model which takes in a celebrity image and recognises that celebrity and objects present in the image. These objects are then used to perform a retrieval on a database of short stories to identify the most thematic story. The named entities are extracted from the seed of the most relevant story, and finally, a customized GPT-2 model takes in this story to generate the resulting short story.

## BACKGROUND

The **VGG Object Detection model** was first introduced in 2014, and it revolutionized the paradigm of object detection in the industry. The model consisted of 16 convolutional layers, and stood out due to its uniform nature. It uses lot of filter layers, and has over 138 million parameters.



Soon, the VGGNet model was extended to perform the task of face recognition, and this new model was called **VGGFace2**. This model was trained on over 9000 celebrity identities and the dataset which it was trained on nearly 3.3 million faces. The faces were chosen after a Google search on the celebrity name. We use the VGGNet object detection model and VGGFace2 dataset to extract objects and faces respectively from a given image.



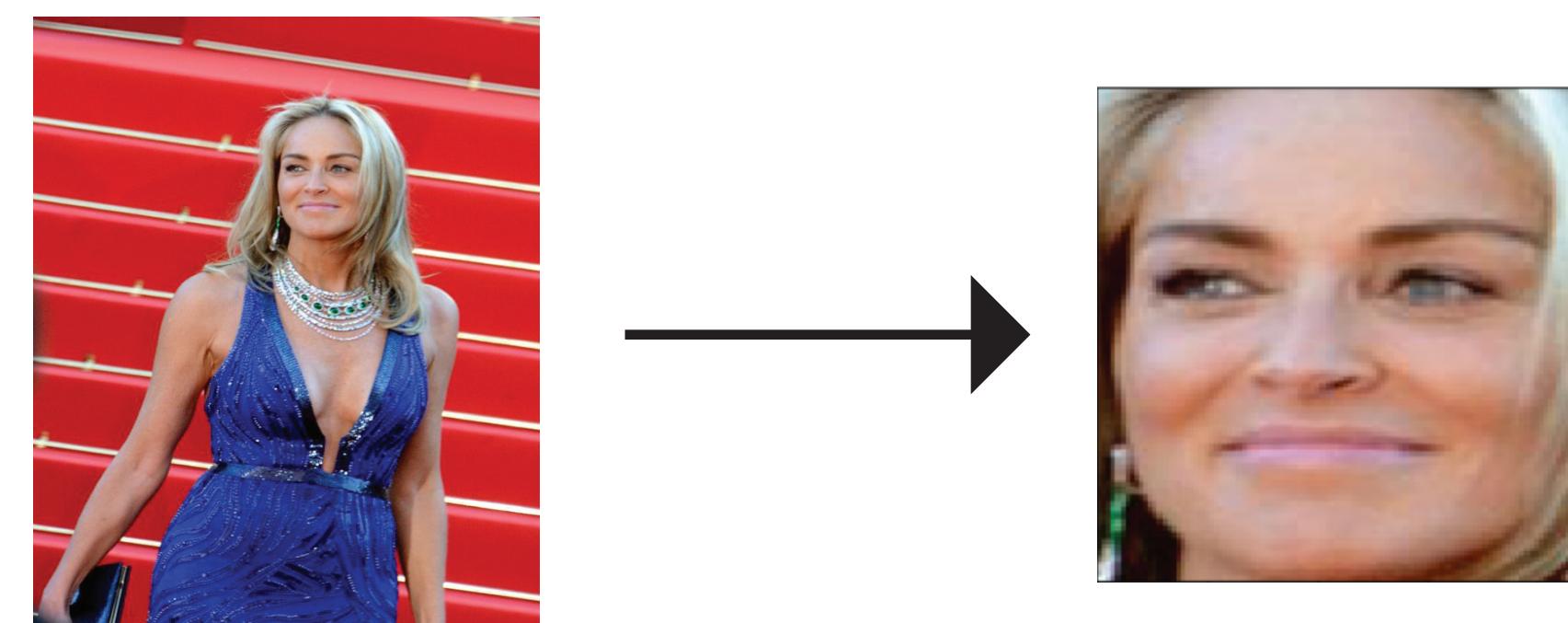
While the most popular technique in the past years has been recurrent neural networks, in particular LSTM networks, Google introduced a new architecture called the **Transformer network** in their ground-breaking paper "Attention is all you need", which stores the context of the text in the network.

OpenAI then used this transformer architecture to create the **GPT-2 model**, game-changing model for text generation which has over 1.5 billion parameters and was trained on over 3 million web pages. However, they did not release the full model immediately due to safety concerns. We used the limited model which has about 345 million parameters.

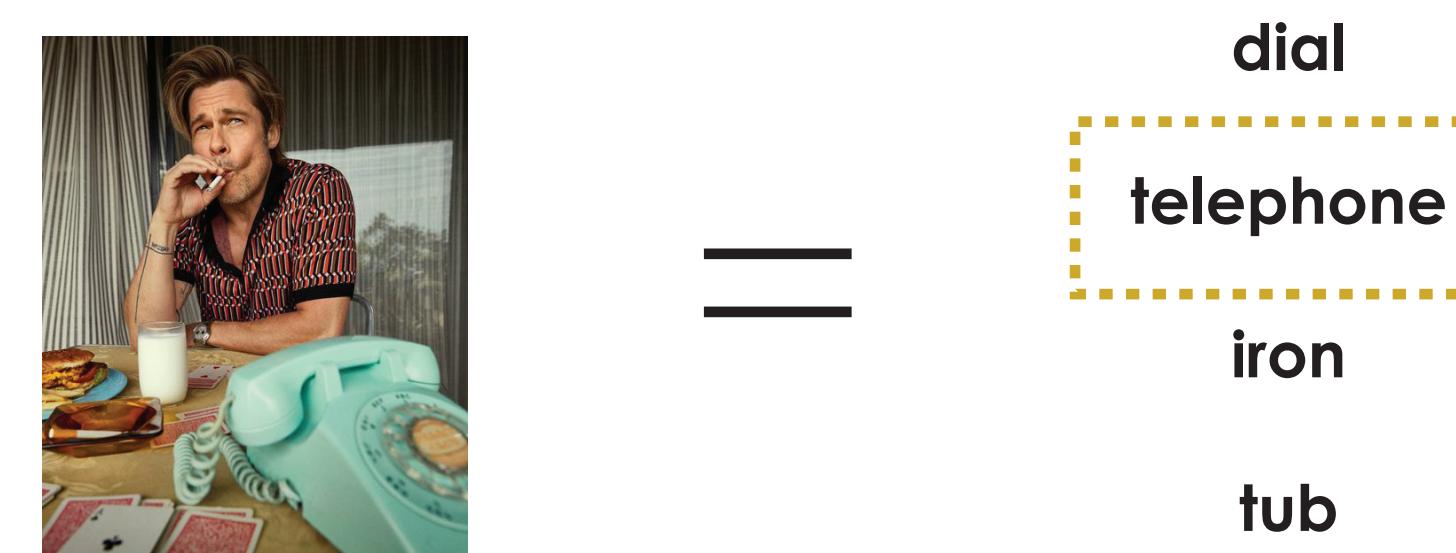
GPT-2 (Generative Pre-trained Transformer) generates text samples in response to the model being primed with an arbitrary input. Fine-tuning this model on small datasets generates near-human like text.

## METHODOLOGY

Given a celebrity image, we first extract the face from the image and then apply pre-trained VGGFace 2 model for classification. It returns the list of possible celebrities identified in the the celebrity image through **Face recognition algorithm** along with probabilities.

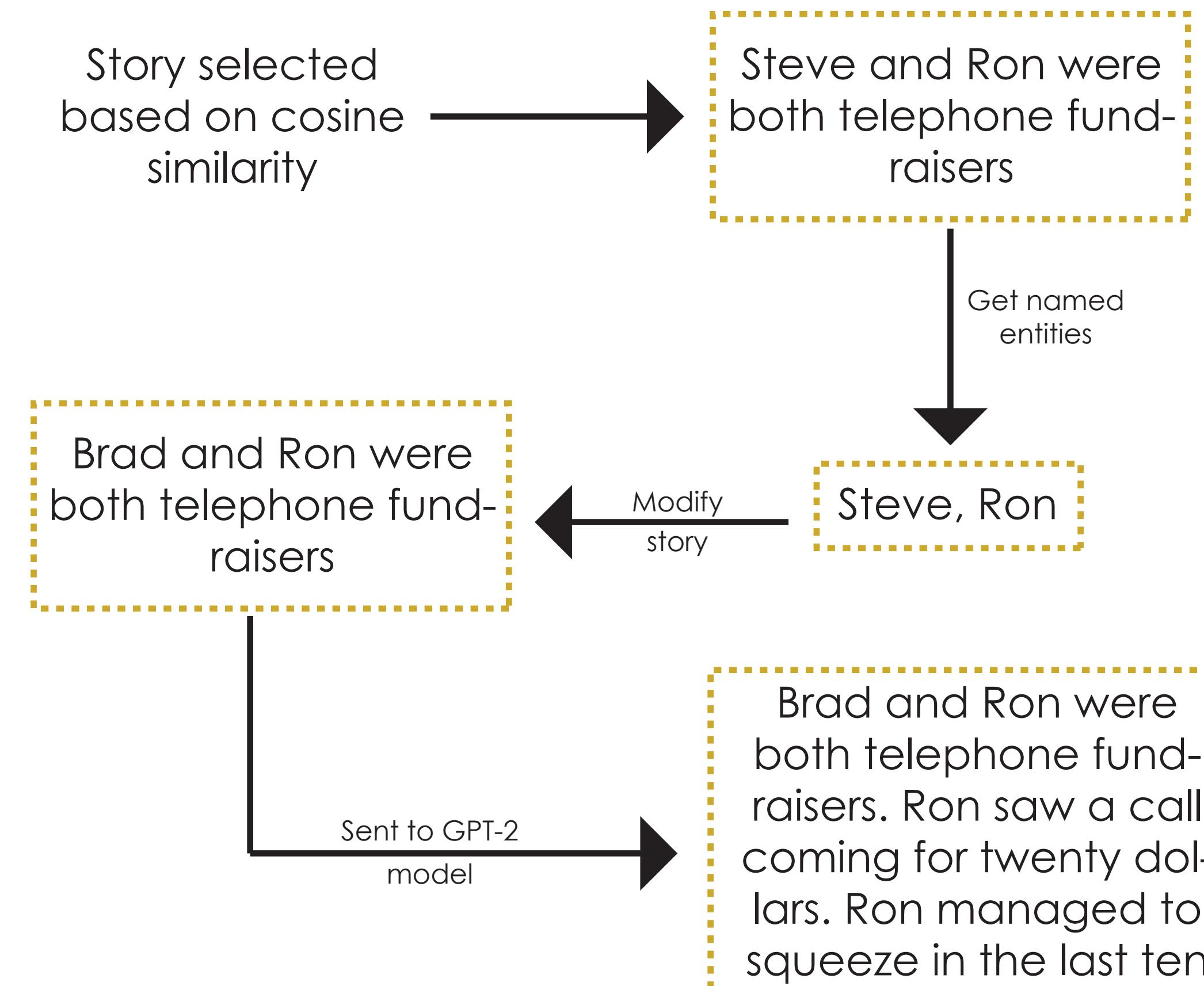


Next, we identified the objects in the celebrity images using VGGNet neural network. It returns the list of possible objects in the image. Below, we see image of Brad Pitt with a telephone and the list of **objects detected** were 'dial', 'telephone', 'iron' and 'tub'.

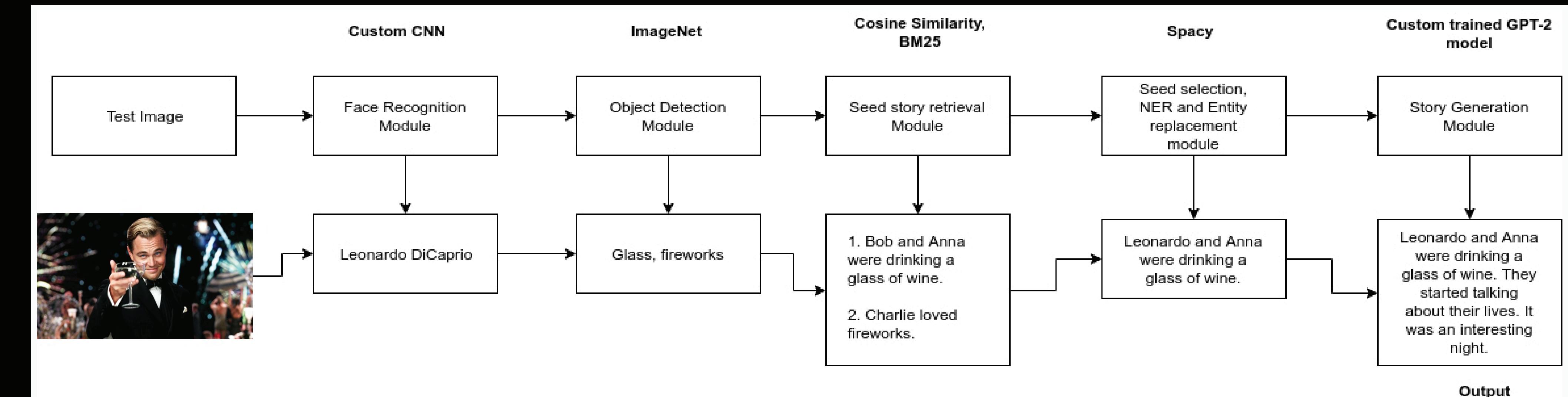


Using the object list generated above, we calculate the cosine similarities for each of the short stories present in our ROC short stories dataset. Further, we keep the story with highest cosine similarity and feed that into Spacy module for **Named Entity recognition**.

We fine-tuned the GPT-2 model by training it on our ROC short stories dataset mentioned above. So the modified story is then sent to the finetuned GPT-2 model to get the generated story.



## ALGORITHM



## GENERATED STORIES



	Katie Perry (99.74%)
	brassiere, bikini, miniskirt, drumstick
	Jane decided to get swimsuit ready
	Jane

Katie decided to get swimsuit ready. She packed everything however, there was one piece missing. She checked and realized that the piece was missing from the tank.



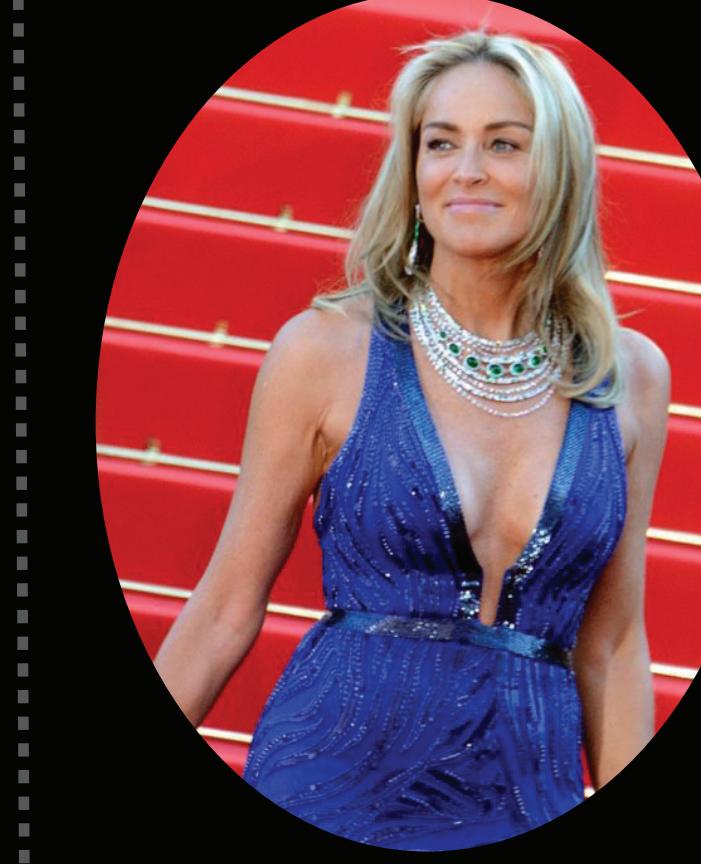
	Brad Pitt (99.31%)
	dial, telephone, iron, tub
	Steve and Ron were both telephone fundraisers
	Steve, Ron

Brad and Ron were both telephone fundraisers. Ron saw a call coming for twenty dollars. Ron managed to squeeze in the last ten donations without concern for whether Ron could do that.



	Channing Tatum (94.44%)
	chain, mail, microphone, crutch
	Ulrich was trying to adjust his microphone stand.
	Ulrich

Channing was trying to adjust his microphone stand. One foot went sliding by him's ears. He lost control and wasn't able to get the balance right before hit.



	Sharon Stone(99.64%)
	miniskirt, cardigan, gown, velvet
	Joey was getting married and needed a gown.
	Joey

Sharon was getting married and needed a gown. She asked her friend for help. Her friend put some glitter around her neck and brought a dress to sell. Sharon was pleased with the dress.