**Chatbot Report**

**Chatbot Description**

This chatbot is intended to be used to help the user learn more about skyscrapers. The intent is that a user can put in the name of a skyscraper and the bot will tell the user a short description of the skyscraper. It can handle some basic greetings for better interactivity. User can type ‘quit’ to end the conversation. User can type ‘history’ to see their chat history.

**NLP Techniques Used**

* Word Tokenization
* Lemmatization
* Parsing user input

**Implementation Overview**

This chatbot basically takes in a JSON array of intents, that mainly contain patterns, responses, and tags, and uses this to train a neural network to predict what tag the user input is most similar to.

This chatbot starts off by reading from a file containing JSON array of intents. Each intent has its own context (not being used), patterns, responses, and tag. Patterns are possible similar queries expected from the user. Responses are possible replies to those patterns. Tags are a summary of the user’s queries and tags act as an de facto identifier of the patterns and of the JSON object itself.

Then, it creates a list of patterns (features) and a list of tags (target) that will be used to train the model. At the same time, using a nested loop and tokenization, a vocabulary of words are extracted from the patterns and a vocabulary of classes (unique tags) are also created here. The vocabulary of words are then lemmatized and any punctuations are removed. Then, the vocabulary of words and classes are sorted and duplicates are removed.

Then, the list of patterns and tags are converted into machine-friendly 1s and 0s using the bag-of-words information retrieval technique. The list of words and classes act as a vocabulary for the patterns and tags. For each pattern and each word in pattern, if the word is in the list of words, it will marked as 1, otherwise, it will be marked as 0. This list of words is put into a bag of word list for the patterns, and another list of 1s is created for the tags. These 2 lists are then converted to training x and y arrays (datasets).

The neutral network is created to train a model. The train x (patterns) and y (tags) datasets get plugged into this model. Model is trained 150 times. The model is trained to determine what a tag a given user input is associated with.

This chatbot takes in a user’s input and passes in the input to a tag prediction function. In the tag prediction function, the input gets tokenized, lemmatized, and converted into array of 1s and 0s based on vocabulary of words. The array gets passed in to the trained model to get a list of probabilities (each containing a probability of tag[n]). Low probability tags get eliminated and remaining list of probabilities gets sorted in descending order. List of probabilities are converted to list of tags. To provide a response to the user, the first tag in the tag list gets picked and the responses for that tag get randomized and one of the responses get outputted to the user.

**Sample Chatbot Dialog**

What is your name? Jibin

I am a chatbot that helps you understand skyscrapers. Type quit if you want to end this conversation. Type history if you want to see your chat history.

Hi!

Hi!

How are you?

I am good.

How are you?

I am okay

Empire State Building

The Empire State Building is a 102-story Art Deco skyscraper in Midtown Manhattan, New York City. The building was designed by Shreve, Lamb & Harmon and built from 1930 to 1931. Its name is derived from "Empire State", the nickname of the state of New York. The building has a roof height of 1,250 feet (380 m) and stands a total of 1,454 feet (443.2 m) tall, including its antenna. The Empire State Building was the world's tallest building until the first tower of the World Trade Center was topped out in 1970; following the September 11 attacks in 2001, the Empire State Building was New York City's tallest building until it was surpassed in 2012. As of 2022, the building is the seventh-tallest building in New York City, the ninth-tallest completed skyscraper in the United States, the 54th-tallest in the world, and the sixth-tallest freestanding structure in the Americas.

The site of the Empire State Building, in Midtown South on the west side of Fifth Avenue between West 33rd and 34th Streets, was developed in 1893 as the Waldorf–Astoria Hotel. In 1929, Empire State Inc. acquired the site and devised plans for a skyscraper there. The design for the Empire State Building was changed fifteen times until it was ensured to be the world's tallest building. Construction started on March 17, 1930, and the building opened thirteen and a half months afterward on May 1, 1931. Despite favorable publicity related to the building's construction, because of the Great Depression and World War II, its owners did not make a profit until the early 1950s.

The building's Art Deco architecture, height, and observation decks have made it a popular attraction. Around four million tourists from around the world annually visit the building's 86th- and 102nd-floor observatories; an additional indoor observatory on the 80th floor opened in 2019. The Empire State Building is an international cultural icon: it has been featured in more than 250 television series and films since the film King Kong was released in 1933. The building's size has become the global standard of reference to describe the height and length of other structures. A symbol of New York City, the building has been named as one of the Seven Wonders of the Modern World by the American Society of Civil Engineers. It was ranked first on the American Institute of Architects' List of America's Favorite Architecture in 2007. Additionally, the Empire State Building and its ground-floor interior were designated city landmarks by the New York City Landmarks Preservation Commission in 1980, and were added to the National Register of Historic Places as a National Historic Landmark in 1986.

Give me a synopsis of the Shanghai Financial Center?

The Shanghai World Financial Center (SWFC; Chinese: 上海环球金融中心; pinyin: Shànghǎi Huánqiú Jīnróng Zhōngxīn, Shanghainese: Zånhae Guejieu Cinyon Tsonsin) is a supertall skyscraper located in the Pudong district of Shanghai. It was designed by Kohn Pedersen Fox and developed by the Mori Building Company, with Leslie E. Robertson Associates as its structural engineer and China State Construction Engineering Corp and Shanghai Construction (Group) General Co. as its main contractor. It is a mixed-use skyscraper, consisting of offices, hotels, conference rooms, observation decks, and ground-floor shopping malls. Park Hyatt Shanghai is the tower's hotel component, comprising 174 rooms and suites occupying the 79th to the 93rd floors, which at the time of completion was the highest hotel in the world. It is now the third-highest hotel in the world after the Ritz-Carlton, Hong Kong, which occupies floors 102 to 118 of the International Commerce Centre.On 14 September 2007, the skyscraper was topped out at 492 meters (1,614.2 ft), making it the 2nd tallest building in the world on completion (the tallest at the time being Taipei 101), the tallest building in the world by roof height only, and the tallest in China. The SWFC opened to the public on 28 August 2008, with its observation deck opening on 30 August. The observation deck offers views from 474 m (1,555 ft) above ground level.The SWFC has been lauded for its design, and in 2008 it was named by architects as the year's best-completed skyscraper. In 2013, the SWFC was exceeded in height by the adjacent Shanghai Tower, which is China's tallest structure as of 2023. Together, The Shanghai World Financial Center, The Shanghai Tower and The Jin Mao Tower form the world's first adjacent grouping of three supertall skyscrapers.

Thanks for the info!

Sorry! I do not understand.

Thanks!

Sorry! I do not understand.

thanks

Sorry! I do not understand.

Thank you

No problem!

history

Bot > I am a chatbot that helps you understand skyscrapers. Type quit if you want to end this conversation. Type history if you want to see your chat history.

Jibin > Hi!

Bot > Hi!

Jibin > How are you?

Bot > I am good.

Jibin > How are you?

Bot > I am okay

Jibin > Empire State Building

Bot > The Empire State Building is a 102-story Art Deco skyscraper in Midtown Manhattan, New York City. The building was designed by Shreve, Lamb & Harmon and built from 1930 to 1931. Its name is derived from "Empire State", the nickname of the state of New York. The building has a roof height of 1,250 feet (380 m) and stands a total of 1,454 feet (443.2 m) tall, including its antenna. The Empire State Building was the world's tallest building until the first tower of the World Trade Center was topped out in 1970; following the September 11 attacks in 2001, the Empire State Building was New York City's tallest building until it was surpassed in 2012. As of 2022, the building is the seventh-tallest building in New York City, the ninth-tallest completed skyscraper in the United States, the 54th-tallest in the world, and the sixth-tallest freestanding structure in the Americas.

The site of the Empire State Building, in Midtown South on the west side of Fifth Avenue between West 33rd and 34th Streets, was developed in 1893 as the Waldorf–Astoria Hotel. In 1929, Empire State Inc. acquired the site and devised plans for a skyscraper there. The design for the Empire State Building was changed fifteen times until it was ensured to be the world's tallest building. Construction started on March 17, 1930, and the building opened thirteen and a half months afterward on May 1, 1931. Despite favorable publicity related to the building's construction, because of the Great Depression and World War II, its owners did not make a profit until the early 1950s.

The building's Art Deco architecture, height, and observation decks have made it a popular attraction. Around four million tourists from around the world annually visit the building's 86th- and 102nd-floor observatories; an additional indoor observatory on the 80th floor opened in 2019. The Empire State Building is an international cultural icon: it has been featured in more than 250 television series and films since the film King Kong was released in 1933. The building's size has become the global standard of reference to describe the height and length of other structures. A symbol of New York City, the building has been named as one of the Seven Wonders of the Modern World by the American Society of Civil Engineers. It was ranked first on the American Institute of Architects' List of America's Favorite Architecture in 2007. Additionally, the Empire State Building and its ground-floor interior were designated city landmarks by the New York City Landmarks Preservation Commission in 1980, and were added to the National Register of Historic Places as a National Historic Landmark in 1986.

Jibin > Give me a synopsis of the Shanghai Financial Center?

Bot > The Shanghai World Financial Center (SWFC; Chinese: 上海环球金融中心; pinyin: Shànghǎi Huánqiú Jīnróng Zhōngxīn, Shanghainese: Zånhae Guejieu Cinyon Tsonsin) is a supertall skyscraper located in the Pudong district of Shanghai. It was designed by Kohn Pedersen Fox and developed by the Mori Building Company, with Leslie E. Robertson Associates as its structural engineer and China State Construction Engineering Corp and Shanghai Construction (Group) General Co. as its main contractor. It is a mixed-use skyscraper, consisting of offices, hotels, conference rooms, observation decks, and ground-floor shopping malls. Park Hyatt Shanghai is the tower's hotel component, comprising 174 rooms and suites occupying the 79th to the 93rd floors, which at the time of completion was the highest hotel in the world. It is now the third-highest hotel in the world after the Ritz-Carlton, Hong Kong, which occupies floors 102 to 118 of the International Commerce Centre.On 14 September 2007, the skyscraper was topped out at 492 meters (1,614.2 ft), making it the 2nd tallest building in the world on completion (the tallest at the time being Taipei 101), the tallest building in the world by roof height only, and the tallest in China. The SWFC opened to the public on 28 August 2008, with its observation deck opening on 30 August. The observation deck offers views from 474 m (1,555 ft) above ground level.The SWFC has been lauded for its design, and in 2008 it was named by architects as the year's best-completed skyscraper. In 2013, the SWFC was exceeded in height by the adjacent Shanghai Tower, which is China's tallest structure as of 2023. Together, The Shanghai World Financial Center, The Shanghai Tower and The Jin Mao Tower form the world's first adjacent grouping of three supertall skyscrapers.

Jibin > Thanks for the info!

Bot > Sorry! I do not understand.

Jibin > Thanks!

Bot > Sorry! I do not understand.

Jibin > thanks

Bot > Sorry! I do not understand.

Jibin > Thank you

Bot > No problem!

Jibin > history

quit

You have successfully ended our conversation. Thank you for chatting with me.

**Knowledge Base Appendix (Current)**

* Hello
* No answer
* How are you
* Thanks
* Skyscraper
* Empire State Building
* Boerentoren
* Edificio Kavanagh
* UN Secretariat Building
* Seagram Building
* Oxford Tower
* Willis Tower
* Messeturm
* Jin Mao Tower
* Shard London Bridge
* Petronas Twin Towers
* One World Trade Center
* Shanghai Tower
* John Hancock Center
* One Chase Manhattan Plaza
* Citigroup Center
* Bank of China Tower
* One Shell Plaza
* Shanghai World Financial Center
* Taipei 101
* 30 St Mary Axe
* Home Insurance Building

**Sample User Model Appendix**

* JPro1
* JPro2

**Evaluation of Chatbot**

I tested the chatbot objectively using 2 models: JPro1 and JPro2. Out of the 10 interactions I made between the 2 users, 7 interactions had correct responses by chatbot. So, I will give it a 70% rating. The strength of this chatbot are that it can handle queries worded slightly differently, it has little to no issue with the cases (upper/lower case) of the queries, it can handle no answers, the summaries are concise and easy to understand, and it can handle some basic greeting. The weaknesses of this chatbot are that it is as smart as the knowledge base it uses, it can only handle what questions, greetings it can handle are limited, it can get queries wrong even when part of query is in tag from knowledge base, poor handling of the unexpected from limited model training and knowledge base.

**High-Level Diagram of Intended Chatbot Logic**

**Loop arrow indicating transition to same state**

**Loop arrow indicating transition to same state**

**User greets chatbot**

**User ends conversation**

**Bot provides synopsis of building**

**Bot greets user**

**User provides skyscraper name user wants to know about**

**User provides skyscraper name user wants to know about**

**User starts conversation**