:Next Word Prediction Using LSTM:

Introduction

In today's fast-paced digital world, efficient communication is paramount. Whether composing emails, drafting documents, or sending messages, the ability to type quickly and accurately is essential. However, traditional typing methods can be time-consuming and prone to errors, hindering productivity and frustrating users.

Enter Next Word Prediction Systems—a groundbreaking technology designed to revolutionize text input. These systems leverage the power of machine learning and natural language processing to anticipate and suggest the most likely next word or phrase as users type, streamlining the typing process and enhancing user experience.

In this, we will delve into the world of next word prediction systems, exploring their importance, functionality, and potential applications. Join us as we uncover the inner workings of these innovative systems and examine their impact on communication, productivity, and accessibility.

Objective

The objective of this report is to demonstrate the effectiveness of Long Short-Term Memory (LSTM) neural networks in predicting the next word of a movie based on preceding sequences of dialogue or screenplay text. The report aims to evaluate the accuracy and performance of the LSTM model in this predictive task, explore the impact of different hyperparameters and training configurations on model performance, and provide insights into the potential applications and limitations of LSTM-based language prediction in the domain of movie dialogue generation and recommendation systems.

Importing Necessary Libraries

```
import pandas as pd
import tensorflow as tf
import numpy as np
```

```
WARNING:tensorflow:From C:\Users\Malkit Phogat\AppData\Roaming\Python\
Python311\site-packages\keras\src\losses.py:2976: The name
tf.losses.sparse_softmax_cross_entropy is deprecated. Please use
tf.compat.v1.losses.sparse_softmax_cross_entropy instead.
```

Import CSV File

```
df = pd.read_csv('tmdb_5000_movies.csv')
df.head(4)
```

```
      budget                                             genres  \
0  237000000  [{"id": 28, "name": "Action"}, {"id": 12, "nam...
1  300000000  [{"id": 12, "name": "Adventure"}, {"id": 14, "...
2  245000000  [{"id": 28, "name": "Action"}, {"id": 12, "nam...
3  250000000  [{"id": 28, "name": "Action"}, {"id": 80, "nam...

                                       homepage      id  \
0                   http://www.avatarmovie.com/   19995
1  http://disney.go.com/disneypictures/pirates/     285
2   http://www.sonypictures.com/movies/spectre/  206647
3            http://www.thedarkknightrises.com/   49026

                                            keywords original_language
\
0  [{"id": 1463, "name": "culture clash"}, {"id":...                en

1  [{"id": 270, "name": "ocean"}, {"id": 726, "na...                en

2  [{"id": 470, "name": "spy"}, {"id": 818, "name...                en

3  [{"id": 849, "name": "dc comics"}, {"id": 853,...                en


                             original_title  \
0                                    Avatar
1  Pirates of the Caribbean: At World's End
2                                   Spectre
3                     The Dark Knight Rises

                                            overview  popularity  \
0  In the 22nd century, a paraplegic Marine is di...  150.437577
1  Captain Barbossa, long believed to be dead, ha...  139.082615
2  A cryptic message from Bond’s past sends him o...  107.376788
3  Following the death of District Attorney Harve...  112.312950

                                production_companies  \
0  [{"name": "Ingenious Film Partners", "id": 289...
1  [{"name": "Walt Disney Pictures", "id": 2}, {"...
2  [{"name": "Columbia Pictures", "id": 5}, {"nam...
3  [{"name": "Legendary Pictures", "id": 923}, {"...

                                production_countries release_date
revenue  \
0  [{"iso_3166_1": "US", "name": "United States o...   10-12-2009
2787965087
```

```
1  [{"iso_3166_1": "US", "name": "United States o...   19-05-2007
961000000
2  [{"iso_3166_1": "GB", "name": "United Kingdom"...   26-10-2015
880674609
3  [{"iso_3166_1": "US", "name": "United States o...   16-07-2012
1084939099

   runtime                                          spoken_languages
status  \
0    162.0  [{"iso_639_1": "en", "name": "English"}, {"iso...
Released
1    169.0           [{"iso_639_1": "en", "name": "English"}]
Released
2    148.0  [{"iso_639_1": "fr", "name": "Fran\u00e7ais"},...
Released
3    165.0           [{"iso_639_1": "en", "name": "English"}]
Released

                                          tagline  \
0                     Enter the World of Pandora.
1  At the end of the world, the adventure begins.
2                           A Plan No One Escapes
3                                The Legend Ends

                                    title  vote_average  vote_count

0                                  Avatar           7.2       11800

1  Pirates of the Caribbean: At World's End           6.9        4500

2                                 Spectre           6.3        4466

3                  The Dark Knight Rises           7.6        9106
```

Extracting the required data from CSv fileL

```
df = df['original_title']

df

0                                       Avatar
1      Pirates of the Caribbean: At World's End
2                                      Spectre
3                       The Dark Knight Rises
4                                 John Carter
                         ...
4798                              El Mariachi
4799                                Newlyweds
```

```
4800                    Signed, Sealed, Delivered
4801                             Shanghai Calling
4802                            My Date with Drew
Name: original_title, Length: 4803, dtype: object
```

Converting data into a list for easy handling

```
movie_name = df.to_list()

movie_name
```

```
['Avatar',
 "Pirates of the Caribbean: At World's End",
 'Spectre',
 'The Dark Knight Rises',
 'John Carter',
 'Spider-Man 3',
 'Tangled',
 'Avengers: Age of Ultron',
 'Harry Potter and the Half-Blood Prince',
 'Batman v Superman: Dawn of Justice',
 'Superman Returns',
 'Quantum of Solace',
 "Pirates of the Caribbean: Dead Man's Chest",
 'The Lone Ranger',
 'Man of Steel',
 'The Chronicles of Narnia: Prince Caspian',
 'The Avengers',
 'Pirates of the Caribbean: On Stranger Tides',
 'Men in Black 3',
 'The Hobbit: The Battle of the Five Armies',
 'The Amazing Spider-Man',
 'Robin Hood',
 'The Hobbit: The Desolation of Smaug',
 'The Golden Compass',
 'King Kong',
 'Titanic',
 'Captain America: Civil War',
 'Battleship',
 'Jurassic World',
 'Skyfall',
 'Spider-Man 2',
 'Iron Man 3',
 'Alice in Wonderland',
 'X-Men: The Last Stand',
 'Monsters University',
 'Transformers: Revenge of the Fallen',
 'Transformers: Age of Extinction',
```

```
'Oz: The Great and Powerful',
'The Amazing Spider-Man 2',
'TRON: Legacy',
'Cars 2',
'Green Lantern',
'Toy Story 3',
'Terminator Salvation',
'Furious 7',
'World War Z',
'X-Men: Days of Future Past',
'Star Trek Into Darkness',
'Jack the Giant Slayer',
'The Great Gatsby',
'Prince of Persia: The Sands of Time',
'Pacific Rim',
'Transformers: Dark of the Moon',
'Indiana Jones and the Kingdom of the Crystal Skull',
'The Good Dinosaur',
'Brave',
'Star Trek Beyond',
'WALL·E',
'Rush Hour 3',
'2012',
'A Christmas Carol',
'Jupiter Ascending',
'The Legend of Tarzan',
'The Chronicles of Narnia: The Lion, the Witch and the Wardrobe',
'X-Men: Apocalypse',
'The Dark Knight',
'Up',
'Monsters vs Aliens',
'Iron Man',
'Hugo',
'Wild Wild West',
'The Mummy: Tomb of the Dragon Emperor',
'Suicide Squad',
'Evan Almighty',
'Edge of Tomorrow',
'Waterworld',
'G.I. Joe: The Rise of Cobra',
'Inside Out',
'The Jungle Book',
'Iron Man 2',
'Snow White and the Huntsman',
'Maleficent',
'Dawn of the Planet of the Apes',
'The Lovers',
'47 Ronin',
'Captain America: The Winter Soldier',
```

```
 'Shrek Forever After',
 'Tomorrowland',
 'Big Hero 6',
 'Wreck-It Ralph',
 'The Polar Express',
 'Independence Day: Resurgence',
 'How to Train Your Dragon',
 'Terminator 3: Rise of the Machines',
 'Guardians of the Galaxy',
 'Interstellar',
 'Inception',
 'シン・ゴジラ',
 'The Hobbit: An Unexpected Journey',
 'The Fast and the Furious',
 'The Curious Case of Benjamin Button',
 'X-Men: First Class',
 'The Hunger Games: Mockingjay - Part 2',
 "The Sorcerer's Apprentice",
 'Poseidon',
 'Alice Through the Looking Glass',
 'Shrek the Third',
 'Warcraft',
 'Terminator Genisys',
 'The Chronicles of Narnia: The Voyage of the Dawn Treader',
 'Pearl Harbor',
 'Transformers',
 'Alexander',
 'Harry Potter and the Order of the Phoenix',
 'Harry Potter and the Goblet of Fire',
 'Hancock',
 'I Am Legend',
 'Charlie and the Chocolate Factory',
 'Ratatouille',
 'Batman Begins',
 'Madagascar: Escape 2 Africa',
 'Night at the Museum: Battle of the Smithsonian',
 'X-Men Origins: Wolverine',
 'The Matrix Revolutions',
 'Frozen',
 'The Matrix Reloaded',
 'Thor: The Dark World',
 'Mad Max: Fury Road',
 'Angels & Demons',
 'Thor',
 'Bolt',
 'G-Force',
 'Wrath of the Titans',
 'Dark Shadows',
 'Mission: Impossible - Rogue Nation',
```

```
 'The Wolfman',
 'Bee Movie',
 'Kung Fu Panda 2',
 'The Last Airbender',
 'Mission: Impossible III',
 'White House Down',
 'Mars Needs Moms',
 'Flushed Away',
 'Pan',
 'Mr. Peabody & Sherman',
 'Troy',
 "Madagascar 3: Europe's Most Wanted",
 'Die Another Day',
 'Ghostbusters',
 'Armageddon',
 'Men in Black II',
 'Beowulf',
 'Kung Fu Panda 3',
 'Mission: Impossible - Ghost Protocol',
 'Rise of the Guardians',
 'Fun with Dick and Jane',
 'The Last Samurai',
 'Exodus: Gods and Kings',
 'Star Trek',
 'Spider-Man',
 'How to Train Your Dragon 2',
 'Gods of Egypt',
 'Stealth',
 'Watchmen',
 'Lethal Weapon 4',
 'Hulk',
 'G.I. Joe: Retaliation',
 'Sahara',
 'Final Fantasy: The Spirits Within',
 'Captain America: The First Avenger',
 'The World Is Not Enough',
 'Master and Commander: The Far Side of the World',
 'The Twilight Saga: Breaking Dawn - Part 2',
 'Happy Feet Two',
 'The Incredible Hulk',
 'The BFG',
 'The Revenant',
 'Turbo',
 'Rango',
 'Penguins of Madagascar',
 'The Bourne Ultimatum',
 'Kung Fu Panda',
 'Ant-Man',
 'The Hunger Games: Catching Fire',
```

```
 'Home',
 'War of the Worlds',
 'Bad Boys II',
 'Puss in Boots',
 'Salt',
 'Noah',
 'The Adventures of Tintin',
 'Harry Potter and the Prisoner of Azkaban',
 'Australia',
 'After Earth',
 'Dinosaur',
 'Night at the Museum: Secret of the Tomb',
 'Megamind',
 "Harry Potter and the Philosopher's Stone",
 'R.I.P.D.',
 'Pirates of the Caribbean: The Curse of the Black Pearl',
 'The Hunger Games: Mockingjay - Part 1',
 'The Da Vinci Code',
 'Rio 2',
 'X2',
 'Fast Five',
 'Sherlock Holmes: A Game of Shadows',
 'Clash of the Titans',
 'Total Recall',
 'The 13th Warrior',
 'The Bourne Legacy',
 'Batman & Robin',
 'How the Grinch Stole Christmas',
 'The Day After Tomorrow',
 'Mission: Impossible II',
 'The Perfect Storm',
 '4: Rise of the Silver Surfer',
 'Life of Pi',
 'Ghost Rider',
 'Jason Bourne',
 "Charlie's Angels: Full Throttle",
 'Prometheus',
 'Stuart Little 2',
 'Elysium',
 'The Chronicles of Riddick',
 'RoboCop',
 'Speed Racer',
 'How Do You Know',
 'Knight and Day',
 'Oblivion',
 'Star Wars: Episode III - Revenge of the Sith',
 'Star Wars: Episode II - Attack of the Clones',
 'Monsters, Inc.',
 'The Wolverine',
```

```
'Star Wars: Episode I - The Phantom Menace',
'The Croods',
'Astérix aux Jeux Olympiques',
'Windtalkers',
"The Huntsman: Winter's War",
'Teenage Mutant Ninja Turtles',
'Gravity',
"Dante's Peak",
'Teenage Mutant Ninja Turtles: Out of the Shadows',
'Fantastic Four',
'Night at the Museum',
'San Andreas',
'Tomorrow Never Dies',
'The Patriot',
"Ocean's Twelve",
'Mr. & Mrs. Smith',
'Insurgent',
'The Aviator',
"Gulliver's Travels",
'The Green Hornet',
'300: Rise of an Empire',
'The Smurfs',
'Home on the Range',
'Allegiant',
'Real Steel',
'The Smurfs 2',
'Speed 2: Cruise Control',
"Ender's Game",
'Live Free or Die Hard',
'The Lord of the Rings: The Fellowship of the Ring',
'Around the World in 80 Days',
'Ali',
'The Cat in the Hat',
'I, Robot',
'Kingdom of Heaven',
'Stuart Little',
'The Princess and the Frog',
'The Martian',
'The Island',
'Town & Country',
'Gone in Sixty Seconds',
'Gladiator',
'Minority Report',
'Harry Potter and the Chamber of Secrets',
'Casino Royale',
'Planet of the Apes',
'Terminator 2: Judgment Day',
'Public Enemies',
'American Gangster',
```

```
 'True Lies',
 'The Taking of Pelham 1 2 3',
 'Little Fockers',
 'The Other Guys',
 'Eraser',
 'Django Unchained',
 'The Hunchback of Notre Dame',
 "The Emperor's New Groove",
 'The Expendables 2',
 'National Treasure',
 'Eragon',
 'Where the Wild Things Are',
 'Epic',
 'The Tourist',
 'End of Days',
 'Blood Diamond',
 'The Wolf of Wall Street',
 'Batman Forever',
 'Starship Troopers',
 'Cloud Atlas',
 "Legend of the Guardians: The Owls of Ga'Hoole",
 'Catwoman',
 'Hercules',
 'Treasure Planet',
 'Land of the Lost',
 'The Expendables 3',
 'Point Break',
 'Son of the Mask',
 'In the Heart of the Sea',
 'The Adventures of Pluto Nash',
 'Green Zone',
 'The Peanuts Movie',
 'The Spanish Prisoner',
 'The Mummy Returns',
 'Gangs of New York',
 '金陵十三釵',
 "Surf's Up",
 'The Stepford Wives',
 'Black Hawk Down',
 'The Campaign',
 'The Fifth Element',
 'Sex and the City 2',
 'The Road to El Dorado',
 'Ice Age: Continental Drift',
 'Cinderella',
 'The Lovely Bones',
 'Finding Nemo',
 'The Lord of the Rings: The Return of the King',
 'The Lord of the Rings: The Two Towers',
```

```
 'Seventh Son',
 'Lara Croft: Tomb Raider',
 'Transcendence',
 'Jurassic Park III',
 'Rise of the Planet of the Apes',
 'The Spiderwick Chronicles',
 'A Good Day to Die Hard',
 'The Alamo',
 'The Incredibles',
 'Cutthroat Island',
 'Percy Jackson & the Olympians: The Lightning Thief',
 'Men in Black',
 'Toy Story 2',
 'Unstoppable',
 'Rush Hour 2',
 'What Lies Beneath',
 'Cloudy with a Chance of Meatballs',
 'Ice Age: Dawn of the Dinosaurs',
 'The Secret Life of Walter Mitty',
 "Charlie's Angels",
 'The Departed',
 'Mulan',
 'Tropic Thunder',
 'The Girl with the Dragon Tattoo',
 'Die Hard: With a Vengeance',
 'Sherlock Holmes',
 'Ben-Hur',
 'Atlantis: The Lost Empire',
 'Alvin and the Chipmunks: The Road Chip',
 'Valkyrie',
 "You Don't Mess with the Zohan",
 'Pixels',
 'A.I. Artificial Intelligence',
 'The Haunted Mansion',
 'Contact',
 'Hollow Man',
 'The Interpreter',
 'Percy Jackson: Sea of Monsters',
 'Lara Croft Tomb Raider: The Cradle of Life',
 'Now You See Me 2',
 'The Saint',
 'Spy Game',
 'Mission to Mars',
 'Rio',
 'Bicentennial Man',
 'Volcano',
 "The Devil's Own",
 'K-19: The Widowmaker',
 'Conan the Barbarian',
```

```
 'Cinderella Man',
 'The Nutcracker: The Untold Story',
 'Seabiscuit',
 'Twister',
 'Cast Away',
 'Happy Feet',
 'The Bourne Supremacy',
 'Air Force One',
 "Ocean's Eleven",
 'The Three Musketeers',
 'Hotel Transylvania',
 'Enchanted',
 'Safe House',
 '102 Dalmatians',
 'Tower Heist',
 'The Holiday',
 'Enemy of the State',
 "It's Complicated",
 "Ocean's Thirteen",
 'Open Season',
 'Divergent',
 'Enemy at the Gates',
 'The Rundown',
 'Last Action Hero',
 'Memoirs of a Geisha',
 'The Fast and the Furious: Tokyo Drift',
 'Arthur Christmas',
 'Meet Joe Black',
 'Collateral Damage',
 'All That Jazz',
 'Mirror Mirror',
 'Scott Pilgrim vs. the World',
 'The Core',
 'Nutty Professor II: The Klumps',
 'Scooby-Doo',
 'Dredd',
 'Click',
 'Creepshow',
 'Cats & Dogs 2 : The Revenge of Kitty Galore',
 'Jumper',
 'Hellboy II: The Golden Army',
 'Zodiac',
 'The 6th Day',
 'Bruce Almighty',
 'The Expendables',
 'Mission: Impossible',
 'The Hunger Games',
 'The Hangover Part II',
 'Batman Returns',
```

```
 'Over the Hedge',
 'Lilo & Stitch',
 "Charlotte's Web",
 'Deep Impact',
 'RED 2',
 'The Longest Yard',
 'Alvin and the Chipmunks: Chipwrecked',
 'Grown Ups 2',
 'Get Smart',
 "Something's Gotta Give",
 'Shutter Island',
 'Four Christmases',
 'Robots',
 'Face/Off',
 'Bedtime Stories',
 'Road to Perdition',
 'Just Go with It',
 'Con Air',
 'Eagle Eye',
 'Cold Mountain',
 'The Book of Eli',
 'Flubber',
 'The Haunting',
 'Space Jam',
 'The Pink Panther',
 'The Day the Earth Stood Still',
 'Conspiracy Theory',
 'Fury',
 'Six Days Seven Nights',
 'Yogi Bear',
 'Spirit: Stallion of the Cimarron',
 'Zookeeper',
 'Lost in Space',
 'The Manchurian Candidate',
 'Déjà Vu',
 'Hotel Transylvania 2',
 'Fantasia 2000',
 'The Time Machine',
 'Mighty Joe Young',
 'Swordfish',
 'The Legend of Zorro',
 'What Dreams May Come',
 'Little Nicky',
 'The Brothers Grimm',
 'Mars Attacks!',
 'Évolution',
 'The Edge',
 'Surrogates',
 'Thirteen Days',
```

```
 'Daylight',
 'Walking With Dinosaurs',
 'Battlefield Earth',
 'Looney Tunes: Back in Action',
 'Nine',
 'Timeline',
 'The Postman',
 'Babe: Pig in the City',
 'The Last Witch Hunter',
 'Red Planet',
 'Arthur et les Minimoys',
 'Oceans',
 'A Sound of Thunder',
 'Pompeii',
 'Don Gato: El inicio de la pandilla',
 'A Beautiful Mind',
 'The Lion King',
 'Journey 2: The Mysterious Island',
 'Cloudy with a Chance of Meatballs 2',
 'Red Dragon',
 'Hidalgo',
 'Jack and Jill',
 '2 Fast 2 Furious',
 'The Little Prince',
 'The Invasion',
 'The Adventures of Rocky & Bullwinkle',
 'The Secret Life of Pets',
 'The League of Extraordinary Gentlemen',
 'Despicable Me 2',
 'Independence Day',
 'The Lost World: Jurassic Park',
 'Madagascar',
 'Children of Men',
 'X-Men',
 'Wanted',
 'The Rock',
 'Ice Age: The Meltdown',
 '50 First Dates',
 'Hairspray',
 'Exorcist: The Beginning',
 'Inspector Gadget',
 'Now You See Me',
 'Grown Ups',
 'The Terminal',
 'Hotel for Dogs',
 'Vertical Limit',
 "Charlie Wilson's War",
 'Shark Tale',
 'Dreamgirls',
 'Be Cool',
```

```
 'Munich',
 'Tears of the Sun',
 'Killers',
 'The Man from U.N.C.L.E.',
 'Spanglish',
 'Monster House',
 'Bandits',
 'First Knight',
 'Anna and the King',
 'Immortals',
 'Hostage',
 'Titan A.E.',
 'Hollywood Homicide',
 'Soldier',
 'Carriers',
 'Monkeybone',
 'Flight of the Phoenix',
 'Unbreakable',
 'Minions',
 'Sucker Punch',
 'Snake Eyes',
 'Sphere',
 'The Angry Birds Movie',
 "Fool's Gold",
 'Funny People',
 'The Kingdom',
 'Talladega Nights: The Ballad of Ricky Bobby',
 'Dr. Dolittle 2',
 'Braveheart',
 'Jarhead',
 'The Simpsons Movie',
 'The Majestic',
 'Driven',
 'Deux frères',
 'The Village',
 'Doctor Dolittle',
 'Signs',
 'Shrek 2',
 'Cars',
 'Runaway Bride',
 'xXx',
 'The SpongeBob Movie: Sponge Out of Water',
 'Ransom',
 'Inglourious Basterds',
 'Hook',
 'Die Hard 2',
 'S.W.A.T.',
 'Vanilla Sky',
 'Lady in the Water',
```

```
'AVP: Alien vs. Predator',
'Alvin and the Chipmunks: The Squeakquel',
'We Were Soldiers',
'Olympus Has Fallen',
'Star Trek: Insurrection',
'Battle: Los Angeles',
'Big Fish',
'Wolf',
'War Horse',
'The Monuments Men',
'The Abyss',
'Wall Street: Money Never Sleeps',
'Dracula Untold',
'The Siege',
'Stardust',
'Seven Years in Tibet',
'The Dilemma',
'Bad Company',
'Doom',
'I Spy',
'Underworld: Awakening',
'Rock of Ages',
"Hart's War",
'Killer Elite',
'Rollerball',
'Ballistic: Ecks vs. Sever',
'Hard Rain',
'Osmosis Jones',
"Legends of Oz: Dorothy's Return",
'Blackhat',
'Sky Captain and the World of Tomorrow',
'Basic Instinct 2',
'Escape Plan',
'The Legend of Hercules',
'The Sum of All Fears',
'The Twilight Saga: Eclipse',
'The Score',
'Despicable Me',
'Money Train',
'Ted 2',
'Agora',
'Mystery Men',
'Hall Pass',
'The Insider',
'The Finest Hours',
'Body of Lies',
'Dinner for Schmucks',
'Abraham Lincoln: Vampire Hunter',
'Entrapment',
```

```
 'The X Files',
 'The Last Legion',
 'Saving Private Ryan',
 'Need for Speed',
 'What Women Want',
 'Ice Age',
 'Dreamcatcher',
 'Lincoln',
 'The Matrix',
 'Apollo 13',
 'The Santa Clause 2',
 'Les Misérables',
 "You've Got Mail",
 'Step Brothers',
 'The Mask of Zorro',
 'Due Date',
 'Unbroken',
 'Space Cowboys',
 'Cliffhanger',
 'Broken Arrow',
 'The Kid',
 'World Trade Center',
 'Mona Lisa Smile',
 'The Dictator',
 'Eyes Wide Shut',
 'Annie',
 'Focus',
 'This Means War',
 'Blade: Trinity',
 'Red Dawn',
 'Primary Colors',
 'Resident Evil: Retribution',
 'Death Race',
 'The Long Kiss Goodnight',
 'Proof of Life',
 'Zathura: A Space Adventure',
 'Fight Club',
 'We Are Marshall',
 'Hudson Hawk',
 'Lucky Numbers',
 'I, Frankenstein',
 'Oliver Twist',
 'Elektra',
 'Sin City: A Dame to Kill For',
 'Random Hearts',
 'Everest',
 'Perfume: The Story of a Murderer',
 'Austin Powers in Goldmember',
 'Astro Boy',
```

```
 'Jurassic Park',
 'Wyatt Earp',
 'Clear and Present Danger',
 '天將雄師',
 'Little Man',
 'U-571',
 'The American President',
 'The Love Guru',
 '3000 Miles to Graceland',
 'The Hateful Eight',
 'Blades of Glory',
 'Hop',
 '300',
 'Meet the Fockers',
 'Marley & Me',
 'The Green Mile',
 'Wild Hogs',
 'Chicken Little',
 'Gone Girl',
 'The Bourne Identity',
 'GoldenEye',
 "The General's Daughter",
 'The Truman Show',
 'The Prince of Egypt',
 'Daddy Day Care',
 '2 Guns',
 'Cats & Dogs',
 'The Italian Job',
 'Two Weeks Notice',
 'Antz',
 'Couples Retreat',
 'Days of Thunder',
 'Cheaper by the Dozen 2',
 'Maze Runner: The Scorch Trials',
 'Eat Pray Love',
 'The Family Man',
 'RED',
 'Any Given Sunday',
 'The Horse Whisperer',
 'Collateral',
 'The Scorpion King',
 'Ladder 49',
 'Jack Reacher',
 'Deep Blue Sea',
 "Michael Jackson's This Is It",
 'Contagion',
 'Kangaroo Jack',
 'Coraline',
 'The Happening',
```

```
 'Man on Fire',
 'The Shaggy Dog',
 'Starsky & Hutch',
 'Jingle All the Way',
 'Hellboy',
 'A Civil Action',
 'ParaNorman',
 'The Jackal',
 'Paycheck',
 'Up Close & Personal',
 'The Tale of Despereaux',
 'The Tuxedo',
 'Under Siege 2: Dark Territory',
 'Jack Ryan: Shadow Recruit',
 'Joy',
 'London Has Fallen',
 'Alien: Resurrection',
 'Shooter',
 'The Boxtrolls',
 'Practical Magic',
 'The Lego Movie',
 'Miss Congeniality 2: Armed and Fabulous',
 'Reign of Fire',
 'Gangster Squad',
 'Year One',
 'Invictus',
 'State of Play',
 'Duplicity',
 'My Favorite Martian',
 'The Sentinel',
 'Planet 51',
 'Star Trek: Nemesis',
 'Intolerable Cruelty',
 'Trouble with the Curve',
 'Edge of Darkness',
 'The Relic',
 'Analyze That',
 'Righteous Kill',
 'Mercury Rising',
 'The Soloist',
 'The Legend of Bagger Vance',
 'Almost Famous',
 'Garfield: A Tail of Two Kitties',
 'xXx: State of the Union',
 'Priest',
 'Sinbad: Legend of the Seven Seas',
 'Event Horizon',
 'Dragonfly',
 'The Black Dahlia',
```

```
 'Flyboys',
 'The Last Castle',
 'Supernova',
 "Winter's Tale",
 'The Mortal Instruments: City of Bones',
 'Meet Dave',
 'Dark Water',
 'Edtv',
 'Inkheart',
 'The Spirit',
 'Mortdecai',
 'In the Name of the King: A Dungeon Siege Tale',
 'Beyond Borders',
 '西游记之孙悟空三打白骨精',
 'The Great Raid',
 'Deadpool',
 'Holy Man',
 'American Sniper',
 'Goosebumps',
 'Just Like Heaven',
 'The Flintstones in Viva Rock Vegas',
 'Rambo III',
 'Leatherheads',
 'The Ridiculous 6',
 'Did You Hear About the Morgans?',
 'The Internship',
 'Resident Evil: Afterlife',
 'Red Tails',
 "The Devil's Advocate",
 "That's My Boy",
 'DragonHeart',
 'After the Sunset',
 'Ghost Rider: Spirit of Vengeance',
 "Captain Corelli's Mandolin",
 'The Pacifier',
 'Walking Tall',
 'Forrest Gump',
 'Alvin and the Chipmunks',
 'Meet the Parents',
 'Pocahontas',
 'Superman',
 'The Nutty Professor',
 'Hitch',
 'George of the Jungle',
 'American Wedding',
 'Captain Phillips',
 'Date Night',
 'Casper',
 'The Equalizer',
```

```
 'Maid in Manhattan',
 'Crimson Tide',
 'The Pursuit of Happyness',
 'Flightplan',
 'Disclosure',
 'City of Angels',
 'Kill Bill: Vol. 1',
 'Bowfinger',
 'Kill Bill: Vol. 2',
 'Tango & Cash',
 'Death Becomes Her',
 'Shanghai Noon',
 'Executive Decision',
 "Mr. Popper's Penguins",
 'The Forbidden Kingdom',
 'Free Birds',
 'Alien³',
 'Evita',
 'Ronin',
 'The Ghost and the Darkness',
 'Paddington',
 'The Watch',
 'The Hunted',
 'Instinct',
 'Stuck on You',
 'Semi-Pro',
 'The Pirates! In an Adventure with Scientists!',
 'Changeling',
 'Chain Reaction',
 'The Fan',
 'The Phantom of the Opera',
 'Elizabeth: The Golden Age',
 'Æon Flux',
 'Gods and Generals',
 'Turbulence',
 'Imagine That',
 'Muppets Most Wanted',
 'Thunderbirds',
 'Burlesque',
 'Un long dimanche de fiançailles',
 'Lolita',
 'D-Tox',
 'Blade II',
 'Seven Pounds',
 'Bullet to the Head',
 'The Godfather: Part III',
 'Elizabethtown',
 'You, Me and Dupree',
 'Superman II',
```

```
 'Gigli',
 "All the King's Men",
 'Shaft',
 'Anastasia',
 'Moulin Rouge!',
 'Domestic Disturbance',
 'Black Mass',
 'Flags of Our Fathers',
 'Law Abiding Citizen',
 'Grindhouse',
 'Beloved',
 'Lucky You',
 'Catch Me If You Can',
 'Zero Dark Thirty',
 'The Break-Up',
 'Mamma Mia!',
 "Valentine's Day",
 'The Dukes of Hazzard',
 'The Thin Red Line',
 'The Change-Up',
 'Man on the Moon',
 'Casino',
 'From Paris with Love',
 'Bulletproof Monk',
 'Me, Myself & Irene',
 'Barnyard',
 'Deck the Halls',
 'The Twilight Saga: New Moon',
 'Shrek',
 'The Adjustment Bureau',
 'Robin Hood: Prince of Thieves',
 'Jerry Maguire',
 'Ted',
 'As Good as It Gets',
 'Patch Adams',
 'Anchorman 2: The Legend Continues',
 'Mr. Deeds',
 'Super 8',
 'Erin Brockovich',
 'How to Lose a Guy in 10 Days',
 '22 Jump Street',
 'Interview with the Vampire',
 'Yes Man',
 'Central Intelligence',
 'Stepmom',
 "Daddy's Home",
 'Into the Woods',
 'Inside Man',
 'Payback',
```

```
 'Congo',
 'We Bought a Zoo',
 'Knowing',
 'Failure to Launch',
 'The Ring Two',
 'Crazy, Stupid, Love.',
 'Garfield',
 'Christmas with the Kranks',
 'Moneyball',
 'Outbreak',
 'Non-Stop',
 'Race to Witch Mountain',
 'V for Vendetta',
 'Shanghai Knights',
 'Curious George',
 'Herbie Fully Loaded',
 "Don't Say a Word",
 'Hansel & Gretel: Witch Hunters',
 'Unfaithful',
 'I Am Number Four',
 'Syriana',
 '13 Hours: The Secret Soldiers of Benghazi',
 'The Book of Life',
 'Firewall',
 'Absolute Power',
 'G.I. Jane',
 'The Game',
 'Silent Hill',
 'The Replacements',
 'American Reunion',
 'The Negotiator',
 'Into the Storm',
 'Beverly Hills Cop III',
 'Gremlins 2: The New Batch',
 'The Judge',
 'The Peacemaker',
 'Resident Evil: Apocalypse',
 'Bridget Jones: The Edge of Reason',
 'Out of Time',
 'On Deadly Ground',
 'The Adventures of Sharkboy and Lavagirl',
 'The Beach',
 'Raising Helen',
 'Ninja Assassin',
 'For Love of the Game',
 'Striptease',
 'Marmaduke',
 'Hereafter',
 'Murder by Numbers',
```

```
 'Assassins',
 'Hannibal Rising',
 'The Story of Us',
 'The Host',
 'Basic',
 'Blood Work',
 'The International',
 'Escape from L.A.',
 'The Iron Giant',
 'The Life Aquatic with Steve Zissou',
 'Free State of Jones',
 'The Life of David Gale',
 'Man of the House',
 'Run All Night',
 'Eastern Promises',
 'Into the Blue',
 'Joan of Arc',
 'Your Highness',
 'Dream House',
 'Mad City',
 "Baby's Day Out",
 'The Scarlet Letter',
 'Fair Game',
 'Domino',
 'Jade',
 'Gamer',
 'Beautiful Creatures',
 'Death to Smoochy',
 'Zoolander 2',
 'The Big Bounce',
 'What Planet Are You From?',
 ...]
```

Tokenizing the data

This code snippet prepares text data for deep learning models. Initially, the Song_name array is converted to a string type to ensure consistency. Then, a Tokenizer object is created and fitted on the movie_name data. This tokenizer maps each unique word in the text data to a unique integer. Finally, the texts_to_sequences method converts the text data into sequences of integers based on the mapping created by the tokenizer. This sequence data can then be used as input to deep learning models for tasks such as text generation or sentiment analysis.

```
tokenizer = tf.keras.preprocessing.text.Tokenizer()
tokenizer.fit_on_texts(movie_name)
seq = tokenizer.texts_to_sequences(movie_name)

seq[:10]
```

```
[[1564],
 [210, 2, 1, 431, 47, 432, 72],
 [1565],
 [1, 52, 211, 1566],
 [212, 601],
 [213, 8, 21],
 [1567],
 [902, 146, 2, 1568],
 [110, 214, 4, 1, 433, 53, 147],
 [173, 340, 261, 85, 2, 903]]
```

This code snippet is used to display the first 10 sequences of integers generated from the movie_name data using the tokenizer. It allows us to inspect the transformed data and ensure that the tokenization process is working as expected

```
tokenizer.word_index
```

```
{'the': 1,
 'of': 2,
 'a': 3,
 'and': 4,
 'in': 5,
 '2': 6,
 'to': 7,
 'man': 8,
 'i': 9,
 'love': 10,
 'with': 11,
 'on': 12,
 'me': 13,
 'my': 14,
 'movie': 15,
 'you': 16,
 'dead': 17,
 'last': 18,
 'for': 19,
 'from': 20,
 '3': 21,
 'good': 22,
 'big': 23,
 'day': 24,
 'house': 25,
 'ii': 26,
 'american': 27,
 'men': 28,
 'it': 29,
 'story': 30,
 'life': 31,
 'black': 32,
```

```
 'girl': 33,
 'all': 34,
 'out': 35,
 'time': 36,
 'night': 37,
 'is': 38,
 'world': 39,
 'days': 40,
 'star': 41,
 'part': 42,
 'new': 43,
 'de': 44,
 'little': 45,
 'la': 46,
 'at': 47,
 'up': 48,
 'an': 49,
 'die': 50,
 'city': 51,
 'dark': 52,
 'blood': 53,
 'game': 54,
 'one': 55,
 'red': 56,
 'blue': 57,
 'your': 58,
 'iii': 59,
 'kill': 60,
 'war': 61,
 'white': 62,
 'street': 63,
 'lost': 64,
 'christmas': 65,
 'legend': 66,
 'wild': 67,
 'how': 68,
 'road': 69,
 'boys': 70,
 'high': 71,
 'end': 72,
 'down': 73,
 'home': 74,
 'hard': 75,
 'boy': 76,
 'year': 77,
 'like': 78,
 'x': 79,
 'bad': 80,
 'do': 81,
```

```
 'get': 82,
 'back': 83,
 'about': 84,
 'dawn': 85,
 'trek': 86,
 'first': 87,
 'mr': 88,
 'final': 89,
 'two': 90,
 'secret': 91,
 'that': 92,
 'we': 93,
 'no': 94,
 'dragon': 95,
 'planet': 96,
 'mad': 97,
 'earth': 98,
 'never': 99,
 'land': 100,
 'return': 101,
 'spy': 102,
 'les': 103,
 'evil': 104,
 'death': 105,
 'by': 106,
 'le': 107,
 'who': 108,
 'friday': 109,
 'harry': 110,
 'fire': 111,
 'escape': 112,
 '4': 113,
 'four': 114,
 'what': 115,
 'this': 116,
 'dog': 117,
 'run': 118,
 'best': 119,
 'thing': 120,
 'king': 121,
 'america': 122,
 'green': 123,
 'vs': 124,
 'rise': 125,
 'snow': 126,
 'after': 127,
 'ghost': 128,
 'perfect': 129,
 'or': 130,
```

```
'island': 131,
'sea': 132,
'york': 133,
'state': 134,
'over': 135,
'go': 136,
'young': 137,
'tale': 138,
'be': 139,
'eyes': 140,
'water': 141,
'her': 142,
'cop': 143,
'baby': 144,
'gun': 145,
'age': 146,
'prince': 147,
'jones': 148,
'beyond': 149,
'joe': 150,
'inside': 151,
'not': 152,
'adventures': 153,
'd': 154,
'true': 155,
'just': 156,
'space': 157,
'brothers': 158,
'rock': 159,
'money': 160,
'got': 161,
'club': 162,
'family': 163,
'wedding': 164,
'our': 165,
'guy': 166,
'woman': 167,
'du': 168,
'when': 169,
'next': 170,
'girls': 171,
'saw': 172,
'batman': 173,
'monsters': 174,
'great': 175,
'jack': 176,
'moon': 177,
'witch': 178,
'book': 179,
```

```
 'journey': 180,
 'max': 181,
 '1': 182,
 '13th': 183,
 'heaven': 184,
 'other': 185,
 'things': 186,
 'el': 187,
 'ice': 188,
 "don't": 189,
 "devil's": 190,
 'hotel': 191,
 'dogs': 192,
 'e': 193,
 'kid': 194,
 'glory': 195,
 'way': 196,
 'hills': 197,
 'dream': 198,
 'drive': 199,
 'room': 200,
 'child': 201,
 'texas': 202,
 'hot': 203,
 'midnight': 204,
 '3d': 205,
 'diary': 206,
 'school': 207,
 'summer': 208,
 'light': 209,
 'pirates': 210,
 'knight': 211,
 'john': 212,
 'spider': 213,
 'potter': 214,
 'iron': 215,
 'revenge': 216,
 'into': 217,
 'darkness': 218,
 'apes': 219,
 'train': 220,
 'games': 221,
 'angels': 222,
 'mission': 223,
 'code': 224,
 'know': 225,
 'wars': 226,
 'town': 227,
 'are': 228,
```

```
'son': 229,
'heart': 230,
'sex': 231,
'park': 232,
'three': 233,
'meet': 234,
'deep': 235,
'eye': 236,
'seven': 237,
'bride': 238,
's': 239,
'13': 240,
'kiss': 241,
'show': 242,
'michael': 243,
'under': 244,
'miss': 245,
'did': 246,
'super': 247,
'silent': 248,
'destination': 249,
'soul': 250,
'party': 251,
'nightmare': 252,
'walk': 253,
'too': 254,
'romeo': 255,
'call': 256,
'valley': 257,
'before': 258,
'married': 259,
'halloween': 260,
'superman': 261,
'battle': 262,
'five': 263,
'hood': 264,
'captain': 265,
'kingdom': 266,
'rush': 267,
'edge': 268,
'tomorrow': 269,
'mars': 270,
'wanted': 271,
'side': 272,
'happy': 273,
'stone': 274,
'curse': 275,
'jason': 276,
'teenage': 277,
```

```
 'ninja': 278,
 'real': 279,
 'live': 280,
 'princess': 281,
 'country': 282,
 'lies': 283,
 'wolf': 284,
 'now': 285,
 "it's": 286,
 'open': 287,
 'nights': 288,
 'mighty': 289,
 'walking': 290,
 'hunter': 291,
 'u': 292,
 'l': 293,
 'monster': 294,
 'people': 295,
 'lady': 296,
 'vampire': 297,
 'women': 298,
 'step': 299,
 'date': 300,
 'broken': 301,
 'long': 302,
 'adventure': 303,
 'lucky': 304,
 'runner': 305,
 'shadow': 306,
 'london': 307,
 'magic': 308,
 'name': 309,
 'vegas': 310,
 'line': 311,
 'paris': 312,
 '8': 313,
 '10': 314,
 'crazy': 315,
 'hill': 316,
 'beverly': 317,
 'work': 318,
 'fat': 319,
 'river': 320,
 'dangerous': 321,
 'kids': 322,
 'again': 323,
 'living': 324,
 'lake': 325,
 'sweet': 326,
```

```
'devil': 327,
'friends': 328,
'bang': 329,
'take': 330,
'greatest': 331,
'lives': 332,
'goes': 333,
'have': 334,
'top': 335,
'il': 336,
'brown': 337,
'paranormal': 338,
'activity': 339,
'v': 340,
'lone': 341,
'chronicles': 342,
'golden': 343,
'terminator': 344,
'hour': 345,
'west': 346,
'tomb': 347,
'forever': 348,
'fast': 349,
'hunger': 350,
'charlie': 351,
'force': 352,
'impossible': 353,
'away': 354,
'most': 355,
'another': 356,
'dick': 357,
'twilight': 358,
'bourne': 359,
'mutant': 360,
'lord': 361,
'gone': 362,
'where': 363,
'point': 364,
'zone': 365,
'thief': 366,
'saint': 367,
'air': 368,
'holiday': 369,
'cold': 370,
'pink': 371,
'still': 372,
'machine': 373,
'et': 374,
'don': 375,
```

```
'league': 376,
'hollywood': 377,
'funny': 378,
'dr': 379,
'alien': 380,
'soldiers': 381,
'los': 382,
'rain': 383,
'plan': 384,
'hours': 385,
'blade': 386,
'resident': 387,
'race': 388,
'play': 389,
'rising': 390,
'shanghai': 391,
'catch': 392,
'can': 393,
'stop': 394,
'number': 395,
'list': 396,
'along': 397,
'here': 398,
'double': 399,
'scream': 400,
'ride': 401,
'bridge': 402,
'5': 403,
'art': 404,
'sunshine': 405,
'truth': 406,
'burn': 407,
'ghosts': 408,
'she': 409,
'there': 410,
'elm': 411,
'40': 412,
'old': 413,
'queen': 414,
'will': 415,
'9': 416,
'without': 417,
'30': 418,
'running': 419,
'der': 420,
'cry': 421,
'welcome': 422,
'brother': 423,
'nothing': 424,
```

```
 "she's": 425,
 'lights': 426,
 'kind': 427,
 'butterfly': 428,
 'being': 429,
 'they': 430,
 'caribbean': 431,
 "world's": 432,
 'half': 433,
 'returns': 434,
 "man's": 435,
 'steel': 436,
 'stranger': 437,
 'robin': 438,
 'jurassic': 439,
 'transformers': 440,
 'oz': 441,
 'furious': 442,
 'z': 443,
 'future': 444,
 'past': 445,
 'lion': 446,
 'aliens': 447,
 'g': 448,
 'jungle': 449,
 'soldier': 450,
 'shrek': 451,
 'hero': 452,
 'glass': 453,
 'am': 454,
 'chocolate': 455,
 'madagascar': 456,
 'frozen': 457,
 'wrath': 458,
 'shadows': 459,
 'kung': 460,
 'jane': 461,
 'gods': 462,
 'kings': 463,
 'master': 464,
 'rider': 465,
 'speed': 466,
 'inc': 467,
 'turtles': 468,
 'mrs': 469,
 'empire': 470,
 'free': 471,
 'rings': 472,
 'groove': 473,
```

```
'national': 474,
'mask': 475,
'lovely': 476,
'vengeance': 477,
'alvin': 478,
'chipmunks': 479,
'haunted': 480,
'see': 481,
'own': 482,
'safe': 483,
'thirteen': 484,
'action': 485,
'stories': 486,
'haunting': 487,
'dreams': 488,
'nine': 489,
'sound': 490,
'beautiful': 491,
'mind': 492,
'beginning': 493,
'angry': 494,
'ballad': 495,
'bobby': 496,
't': 497,
'sky': 498,
'predator': 499,
'fish': 500,
'horse': 501,
'dracula': 502,
'years': 503,
'company': 504,
'underworld': 505,
'hall': 506,
'body': 507,
'saving': 508,
'want': 509,
'santa': 510,
'trade': 511,
'sin': 512,
'hearts': 513,
'chicken': 514,
'daddy': 515,
'job': 516,
'close': 517,
'joy': 518,
'trouble': 519,
'george': 520,
'bill': 521,
'chain': 522,
'un': 523,
```

```
 'head': 524,
 'godfather': 525,
 'zero': 526,
 'as': 527,
 'say': 528,
 'us': 529,
 'fair': 530,
 'sight': 531,
 'riding': 532,
 'liar': 533,
 'scary': 534,
 'hell': 535,
 'tin': 536,
 'johnny': 537,
 'st': 538,
 '21': 539,
 'came': 540,
 'minutes': 541,
 'dance': 542,
 'honor': 543,
 'pitch': 544,
 'november': 545,
 'mean': 546,
 'streets': 547,
 'something': 548,
 'wife': 549,
 'so': 550,
 'virgin': 551,
 'fear': 552,
 'radio': 553,
 'must': 554,
 'business': 555,
 'freedom': 556,
 'history': 557,
 'letters': 558,
"what's": 559,
 'chasing': 560,
 'once': 561,
 'grace': 562,
 'straight': 563,
 'only': 564,
 'luck': 565,
 'sisterhood': 566,
 'right': 567,
 'blonde': 568,
 'crocodile': 569,
 'dundee': 570,
 'ever': 571,
 'but': 572,
```

```
'god': 573,
'omega': 574,
'hit': 575,
'alone': 576,
'jackass': 577,
'let': 578,
'rules': 579,
'dancer': 580,
'das': 581,
'enter': 582,
'everything': 583,
'nowhere': 584,
'born': 585,
'valentine': 586,
'brooklyn': 587,
'dirty': 588,
'pet': 589,
'b': 590,
'musical': 591,
'chapter': 592,
'low': 593,
'better': 594,
'wind': 595,
'saints': 596,
'wicked': 597,
'des': 598,
'20': 599,
'ha': 600,
'carter': 601,
'narnia': 602,
'hobbit': 603,
'amazing': 604,
'titanic': 605,
'civil': 606,
'alice': 607,
'wonderland': 608,
'stand': 609,
'fallen': 610,
'cars': 611,
'toy': 612,
'indiana': 613,
'brave': 614,
'carol': 615,
'apocalypse': 616,
'lovers': 617,
'6': 618,
'express': 619,
'independence': 620,
'guardians': 621,
```

```
 'galaxy': 622,
 'through': 623,
 'looking': 624,
 'voyage': 625,
 'pearl': 626,
 'factory': 627,
 'africa': 628,
 'museum': 629,
 'matrix': 630,
 'titans': 631,
 'nation': 632,
 'fu': 633,
 'panda': 634,
 'samurai': 635,
 'weapon': 636,
 'avenger': 637,
 'enough': 638,
 'far': 639,
 'saga': 640,
 'feet': 641,
 'penguins': 642,
 'prisoner': 643,
 'r': 644,
 'sherlock': 645,
 'holmes': 646,
 'warrior': 647,
 'storm': 648,
 'surfer': 649,
 'full': 650,
 'episode': 651,
 'phantom': 652,
 "winter's": 653,
 "ocean's": 654,
 'smith': 655,
 'range': 656,
 'control': 657,
 'cat': 658,
 'hat': 659,
 'martian': 660,
 'secrets': 661,
 'casino': 662,
 'expendables': 663,
 'wall': 664,
 'atlas': 665,
 'break': 666,
 'cinderella': 667,
 'finding': 668,
 'jackson': 669,
 'beneath': 670,
```

```
'thunder': 671,
'cradle': 672,
'k': 673,
'heist': 674,
'enemy': 675,
'mirror': 676,
'cats': 677,
'impact': 678,
'longest': 679,
'give': 680,
'face': 681,
'off': 682,
'mountain': 683,
'conspiracy': 684,
'bear': 685,
'spirit': 686,
'2000': 687,
'may': 688,
'come': 689,
'rocky': 690,
'extraordinary': 691,
'50': 692,
'exorcist': 693,
'shark': 694,
'cool': 695,
'sun': 696,
'killers': 697,
'n': 698,
'anna': 699,
'flight': 700,
'punch': 701,
'gold': 702,
'doctor': 703,
'were': 704,
'has': 705,
'siege': 706,
'killer': 707,
'legends': 708,
'mystery': 709,
'lincoln': 710,
'private': 711,
'due': 712,
'center': 713,
'annie': 714,
'fight': 715,
'marshall': 716,
'frankenstein': 717,
'austin': 718,
'powers': 719,
```

```
 'eight': 720,
 'mile': 721,
 'identity': 722,
 'daughter': 723,
 'weeks': 724,
 'maze': 725,
 'sunday': 726,
 'resurrection': 727,
 '51': 728,
 'mortal': 729,
 'holy': 730,
 'tall': 731,
 'pursuit': 732,
 'vol': 733,
 'tango': 734,
 'elizabeth': 735,
 'muppets': 736,
 "king's": 737,
 'law': 738,
 'lose': 739,
 'interview': 740,
 'loaded': 741,
 'ground': 742,
 'beach': 743,
 'raising': 744,
 'steve': 745,
 'letter': 746,
 'fighter': 747,
 'boat': 748,
 'bone': 749,
 'taken': 750,
 'cable': 751,
 'mary': 752,
 'diaries': 753,
 'royal': 754,
 'quest': 755,
 'english': 756,
 'pretty': 757,
 'miracle': 758,
 'bounty': 759,
 'gate': 760,
 '28': 761,
 'sense': 762,
 'fifty': 763,
 'shades': 764,
 'their': 765,
 'network': 766,
 'julia': 767,
 'dumb': 768,
```

```
 'below': 769,
 'greek': 770,
 'small': 771,
 'horrible': 772,
 'skeleton': 773,
 'made': 774,
 'music': 775,
 'paul': 776,
 'trip': 777,
 'million': 778,
 'morning': 779,
 'extreme': 780,
 'words': 781,
 'grudge': 782,
 'match': 783,
 'single': 784,
 'dolphin': 785,
 "we're": 786,
 'dragons': 787,
 'sorority': 788,
 'banks': 789,
 'lions': 790,
 'femme': 791,
 'desert': 792,
 'father': 793,
 'alive': 794,
 'transporter': 795,
 'attraction': 796,
 'repo': 797,
 'vacation': 798,
 'remember': 799,
 'hunt': 800,
 'calls': 801,
 'dollar': 802,
 'rugrats': 803,
 'fall': 804,
 'hope': 805,
 'sarah': 806,
 'bottle': 807,
 'mike': 808,
 'naked': 809,
 'rabbit': 810,
 'yards': 811,
 'pride': 812,
 'cave': 813,
 'united': 814,
 'faith': 815,
 'woodstock': 816,
 'kick': 817,
```

```
 'very': 818,
 'prejudice': 819,
 'door': 820,
 'zombies': 821,
 'vi': 822,
 'blues': 823,
 'noise': 824,
 'happened': 825,
 'o': 826,
 'agent': 827,
 'cody': 828,
 'highlander': 829,
 'jimmy': 830,
 'genius': 831,
 'bunny': 832,
 'beauty': 833,
 'grave': 834,
 'august': 835,
 'tora': 836,
 'dear': 837,
 'bright': 838,
 'lords': 839,
 'iv': 840,
 'think': 841,
 'fisher': 842,
 'march': 843,
 'jaws': 844,
 'animal': 845,
 'wimpy': 846,
 'funeral': 847,
 'molly': 848,
 'song': 849,
 '12': 850,
 'hole': 851,
 'police': 852,
 'chainsaw': 853,
 'harold': 854,
 'kumar': 855,
 'psycho': 856,
 'kills': 857,
 "can't": 858,
 'horror': 859,
 'las': 860,
 'exorcism': 861,
 'rose': 862,
 'happens': 863,
 'station': 864,
 'anything': 865,
 "let's": 866,
```

```
'tales': 867,
'massacre': 868,
'teen': 869,
'cut': 870,
'place': 871,
'voice': 872,
'ones': 873,
'dude': 874,
'car': 875,
'effect': 876,
'project': 877,
'always': 878,
'purge': 879,
'insidious': 880,
'sugar': 881,
'heroes': 882,
'film': 883,
'hai': 884,
' ': 885,
'000': 886,
'food': 887,
'falls': 888,
'kevin': 889,
'mother': 890,
'h': 891,
'east': 892,
'myers': 893,
'was': 894,
'shine': 895,
'lane': 896,
'billy': 897,
'together': 898,
'times': 899,
'giants': 900,
'roadside': 901,
'avengers': 902,
'justice': 903,
'ranger': 904,
'tides': 905,
'battleship': 906,
'extinction': 907,
'legacy': 908,
'salvation': 909,
'7': 910,
'giant': 911,
'slayer': 912,
'sands': 913,
'dinosaur': 914,
'mummy': 915,
```

```
 'emperor': 916,
 'suicide': 917,
 'squad': 918,
 'almighty': 919,
 'huntsman': 920,
 'ronin': 921,
 'winter': 922,
 'ralph': 923,
 'curious': 924,
 'case': 925,
 'benjamin': 926,
 'class': 927,
 'mockingjay': 928,
 'third': 929,
 'alexander': 930,
 'order': 931,
 'phoenix': 932,
 'begins': 933,
 'origins': 934,
 'wolverine': 935,
 'thor': 936,
 'fury': 937,
 'bee': 938,
 'moms': 939,
 'armageddon': 940,
 'fun': 941,
 'egypt': 942,
 'lethal': 943,
 'hulk': 944,
 'within': 945,
 'commander': 946,
 'breaking': 947,
 'incredible': 948,
 'ant': 949,
 'worlds': 950,
 'p': 951,
 'da': 952,
 'rio': 953,
 'silver': 954,
 'pi': 955,
 "charlie's": 956,
 'stuart': 957,
 'riddick': 958,
 'robocop': 959,
 'attack': 960,
 'menace': 961,
 'gravity': 962,
 'fantastic': 963,
 'san': 964,
```

```
 'patriot': 965,
 '300': 966,
 'smurfs': 967,
 'ring': 968,
 'robot': 969,
 'report': 970,
 'judgment': 971,
 'gangster': 972,
 'taking': 973,
 'fockers': 974,
 'guys': 975,
 'dame': 976,
 "emperor's": 977,
 'treasure': 978,
 'epic': 979,
 'diamond': 980,
 'troopers': 981,
 'cloud': 982,
 'hercules': 983,
 'wives': 984,
 'hawk': 985,
 'fifth': 986,
 'drift': 987,
 'bones': 988,
 'lara': 989,
 'croft': 990,
 'raider': 991,
 'percy': 992,
 'cloudy': 993,
 'chance': 994,
 'meatballs': 995,
 'dinosaurs': 996,
 'walter': 997,
 'tattoo': 998,
 'atlantis': 999,
 'intelligence': 1000,
 ...}
```

Removing the movies with a single word as we cant predict the next word of them:

```
X = []
y = []
total_words_dropped = 0

for i in seq:
    if len(i) > 1:
```

```
        for index in range(1, len(i)):
            X.append(i[:index])
            y.append(i[index])
    else:
        total_words_dropped += 1

print("Total Single Words Dropped are:", total_words_dropped)
```

```
Total Single Words Dropped are: 1003
```

```
X[:10]
```

```
[[210],
 [210, 2],
 [210, 2, 1],
 [210, 2, 1, 431],
 [210, 2, 1, 431, 47],
 [210, 2, 1, 431, 47, 432],
 [1],
 [1, 52],
 [1, 52, 211],
 [212]]
```

```
y[:10]
```

```
[2, 1, 431, 47, 432, 72, 52, 211, 1566, 601]
```

Adding Sequence padding to ensure that input sequences have the same length by padding them with a specific value

This line of code uses the pad_sequences function from the Keras preprocessing module to pad the input sequences X to ensure that they all have the same length. This is necessary for training neural networks where inputs need to have uniform dimensions

```
X = tf.keras.preprocessing.sequence.pad_sequences(X)
```

```
X
```

```
array([[  0,   0,   0, ...,   0,   0, 210],
       [  0,   0,   0, ...,   0, 210,   2],
       [  0,   0,   0, ..., 210,   2,   1],
       ...,
       [  0,   0,   0, ...,   0,   0,  14],
       [  0,   0,   0, ...,   0,  14, 300],
       [  0,   0,   0, ...,  14, 300,  11]])
```

```
X.shape
```

```
(8483, 14)
```

```
y = tf.keras.utils.to_categorical(y)
```

```
y
```

```
array([[0., 0., 1., ..., 0., 0., 0.],
       [0., 1., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       ...,
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.],
       [0., 0., 0., ..., 0., 0., 0.]], dtype=float32)
```

```
y.shape
```

```
(8483, 5046)
```

Calculating the total number of words in our vocabulary

This line of code calculates the size of the vocabulary by adding 1 to the number of unique words in the tokenizer.word_index. The additional 1 is for the padding token.

```
vocab_size = len(tokenizer.word_index) + 1 # Calculate The Size Of Vocabulary
```

```
vocab_size
```

```
5046
```

Creating the model

```
model = tf.keras.Sequential([
    tf.keras.layers.Embedding(vocab_size, 14),
    tf.keras.layers.LSTM(100, return_sequences=True),
    tf.keras.layers.LSTM(100),
    tf.keras.layers.Dense(100, activation='relu'),
    tf.keras.layers.Dense(vocab_size, activation='softmax'),
])
```

```
WARNING:tensorflow:From C:\Users\Malkit Phogat\AppData\Roaming\Python\Python311\site-packages\keras\src\backend.py:873: The name tf.get_default_graph is deprecated. Please use
```

```
tf.compat.v1.get_default_graph instead.
```

Printing the summary of the model

```
model.summary()
```

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #   
=================================================================
 embedding (Embedding)       (None, None, 14)          70644     

 lstm (LSTM)                 (None, None, 100)         46000     

 lstm_1 (LSTM)               (None, 100)               80400     

 dense (Dense)               (None, 100)               10100     

 dense_1 (Dense)             (None, 5046)              509646    

=================================================================
Total params: 716790 (2.73 MB)
Trainable params: 716790 (2.73 MB)
Non-trainable params: 0 (0.00 Byte)
_________________________________________________________________
```

Here's the explanation:

Embedding Layer (embedding):

Output Shape: (None, None, 14) Number of Parameters: 70,644 This layer converts integer indices into dense vectors of fixed size. The output shape indicates that it takes sequences of any length and converts each word index into a dense vector of size 14. LSTM Layer (lstm):

Output Shape: (None, None, 100) Number of Parameters: 46,000 LSTM layer with 100 units, returns sequences. It processes sequences of word vectors and outputs sequences of hidden states. LSTM Layer (lstm_1):

Output Shape: (None, 100) Number of Parameters: 80,400 LSTM layer with 100 units. It processes sequences of word vectors and outputs the final hidden state. Dense Layer (dense):

Output Shape: (None, 100) Number of Parameters: 10,100 Fully connected dense layer with 100 units and ReLU activation function. Dense Layer (dense_1):

Output Shape: (None, 5046) Number of Parameters: 509646 Fully connected dense layer with 5046 units and softmax activation function. It outputs the probability distribution over the vocabulary. Total trainable parameters: 716790

Non-trainable parameters: 0

Optimising the model using adam optimiser to minimise the loss function

```
model.compile(
    optimizer=tf.keras.optimizers.Adam(learning_rate=0.004),
    loss='categorical_crossentropy',
    metrics=['accuracy'])

model.fit(X, y, epochs=150)
```

```
Epoch 1/150
WARNING:tensorflow:From C:\Users\Malkit Phogat\AppData\Roaming\Python\
Python311\site-packages\keras\src\utils\tf_utils.py:492: The name
tf.ragged.RaggedTensorValue is deprecated. Please use
tf.compat.v1.ragged.RaggedTensorValue instead.

WARNING:tensorflow:From C:\Users\Malkit Phogat\AppData\Roaming\Python\
Python311\site-packages\keras\src\engine\base_layer_utils.py:384: The
name tf.executing_eagerly_outside_functions is deprecated. Please use
tf.compat.v1.executing_eagerly_outside_functions instead.

266/266 [==============================] - 17s 33ms/step - loss:
7.6410 - accuracy: 0.0585
Epoch 2/150
266/266 [==============================] - 9s 33ms/step - loss: 7.0122
- accuracy: 0.0641
Epoch 3/150
266/266 [==============================] - 9s 33ms/step - loss: 6.9023
- accuracy: 0.0758
Epoch 4/150
266/266 [==============================] - 9s 33ms/step - loss: 7.0542
- accuracy: 0.0606
Epoch 5/150
266/266 [==============================] - 9s 33ms/step - loss: 6.8391
- accuracy: 0.0675
Epoch 6/150
266/266 [==============================] - 9s 33ms/step - loss: 6.7531
- accuracy: 0.0780
Epoch 7/150
266/266 [==============================] - 9s 33ms/step - loss: 6.6687
- accuracy: 0.0866
Epoch 8/150
266/266 [==============================] - 9s 33ms/step - loss: 6.5909
- accuracy: 0.0917
Epoch 9/150
266/266 [==============================] - 9s 33ms/step - loss: 6.5223
- accuracy: 0.0962
Epoch 10/150
266/266 [==============================] - 9s 33ms/step - loss: 6.4621
```

```
- accuracy: 0.1044
Epoch 11/150
266/266 [==============================] - 9s 33ms/step - loss: 6.4135
- accuracy: 0.1077
Epoch 12/150
266/266 [==============================] - 9s 33ms/step - loss: 6.3687
- accuracy: 0.1083
Epoch 13/150
266/266 [==============================] - 9s 33ms/step - loss: 6.3235
- accuracy: 0.1086
Epoch 14/150
266/266 [==============================] - 9s 33ms/step - loss: 6.2789
- accuracy: 0.1103
Epoch 15/150
266/266 [==============================] - 9s 33ms/step - loss: 6.2386
- accuracy: 0.1107
Epoch 16/150
266/266 [==============================] - 9s 33ms/step - loss: 6.1991
- accuracy: 0.1110
Epoch 17/150
266/266 [==============================] - 9s 33ms/step - loss: 6.1619
- accuracy: 0.1106
Epoch 18/150
266/266 [==============================] - 9s 33ms/step - loss: 6.1326
- accuracy: 0.1121
Epoch 19/150
266/266 [==============================] - 9s 33ms/step - loss: 6.0988
- accuracy: 0.1115
Epoch 20/150
266/266 [==============================] - 9s 34ms/step - loss: 6.0677
- accuracy: 0.1122
Epoch 21/150
266/266 [==============================] - 9s 33ms/step - loss: 6.0445
- accuracy: 0.1133
Epoch 22/150
266/266 [==============================] - 9s 34ms/step - loss: 6.0187
- accuracy: 0.1120
Epoch 23/150
266/266 [==============================] - 9s 33ms/step - loss: 5.9885
- accuracy: 0.1146
Epoch 24/150
266/266 [==============================] - 9s 34ms/step - loss: 5.9658
- accuracy: 0.1162
Epoch 25/150
266/266 [==============================] - 9s 33ms/step - loss: 5.9411
- accuracy: 0.1161
Epoch 26/150
266/266 [==============================] - 9s 33ms/step - loss: 5.9187
- accuracy: 0.1148
```

```
Epoch 27/150
266/266 [==============================] - 9s 33ms/step - loss: 5.8984
- accuracy: 0.1158
Epoch 28/150
266/266 [==============================] - 9s 33ms/step - loss: 5.8721
- accuracy: 0.1189
Epoch 29/150
266/266 [==============================] - 9s 33ms/step - loss: 5.8509
- accuracy: 0.1166
Epoch 30/150
266/266 [==============================] - 9s 33ms/step - loss: 5.8332
- accuracy: 0.1209
Epoch 31/150
266/266 [==============================] - 9s 32ms/step - loss: 5.8126
- accuracy: 0.1184
Epoch 32/150
266/266 [==============================] - 9s 33ms/step - loss: 5.7943
- accuracy: 0.1198
Epoch 33/150
266/266 [==============================] - 9s 33ms/step - loss: 5.7690
- accuracy: 0.1212
Epoch 34/150
266/266 [==============================] - 9s 33ms/step - loss: 5.7507
- accuracy: 0.1225
Epoch 35/150
266/266 [==============================] - 9s 33ms/step - loss: 5.7243
- accuracy: 0.1232
Epoch 36/150
266/266 [==============================] - 9s 33ms/step - loss: 5.6996
- accuracy: 0.1222
Epoch 37/150
266/266 [==============================] - 9s 33ms/step - loss: 5.6883
- accuracy: 0.1224
Epoch 38/150
266/266 [==============================] - 9s 33ms/step - loss: 5.6657
- accuracy: 0.1234
Epoch 39/150
266/266 [==============================] - 9s 34ms/step - loss: 5.6517
- accuracy: 0.1226
Epoch 40/150
266/266 [==============================] - 9s 33ms/step - loss: 5.6258
- accuracy: 0.1235
Epoch 41/150
266/266 [==============================] - 9s 33ms/step - loss: 5.6035
- accuracy: 0.1264
Epoch 42/150
266/266 [==============================] - 9s 33ms/step - loss: 5.5886
- accuracy: 0.1265
Epoch 43/150
```

```
266/266 [==============================] - 9s 33ms/step - loss: 5.5758
- accuracy: 0.1268
Epoch 44/150
266/266 [==============================] - 9s 33ms/step - loss: 5.5519
- accuracy: 0.1284
Epoch 45/150
266/266 [==============================] - 9s 34ms/step - loss: 5.6106
- accuracy: 0.1267
Epoch 46/150
266/266 [==============================] - 9s 33ms/step - loss: 5.5510
- accuracy: 0.1275
Epoch 47/150
266/266 [==============================] - 9s 33ms/step - loss: 5.5112
- accuracy: 0.1314
Epoch 48/150
266/266 [==============================] - 9s 33ms/step - loss: 5.4805
- accuracy: 0.1318
Epoch 49/150
266/266 [==============================] - 9s 33ms/step - loss: 5.4668
- accuracy: 0.1331
Epoch 50/150
266/266 [==============================] - 9s 33ms/step - loss: 5.4498
- accuracy: 0.1345
Epoch 51/150
266/266 [==============================] - 9s 33ms/step - loss: 5.4297
- accuracy: 0.1324
Epoch 52/150
266/266 [==============================] - 9s 33ms/step - loss: 5.4128
- accuracy: 0.1352
Epoch 53/150
266/266 [==============================] - 9s 33ms/step - loss: 5.3903
- accuracy: 0.1338
Epoch 54/150
266/266 [==============================] - 9s 33ms/step - loss: 5.3746
- accuracy: 0.1357
Epoch 55/150
266/266 [==============================] - 9s 33ms/step - loss: 5.3565
- accuracy: 0.1370
Epoch 56/150
266/266 [==============================] - 9s 33ms/step - loss: 5.3387
- accuracy: 0.1371
Epoch 57/150
266/266 [==============================] - 9s 33ms/step - loss: 5.3223
- accuracy: 0.1382
Epoch 58/150
266/266 [==============================] - 9s 33ms/step - loss: 5.3091
- accuracy: 0.1404
Epoch 59/150
266/266 [==============================] - 9s 33ms/step - loss: 5.2937
```

```
- accuracy: 0.1391
Epoch 60/150
266/266 [==============================] - 9s 33ms/step - loss: 5.2730
- accuracy: 0.1417
Epoch 61/150
266/266 [==============================] - 9s 33ms/step - loss: 5.2617
- accuracy: 0.1413
Epoch 62/150
266/266 [==============================] - 9s 33ms/step - loss: 5.2417
- accuracy: 0.1437
Epoch 63/150
266/266 [==============================] - 9s 33ms/step - loss: 5.2263
- accuracy: 0.1453
Epoch 64/150
266/266 [==============================] - 9s 33ms/step - loss: 5.2165
- accuracy: 0.1443
Epoch 65/150
266/266 [==============================] - 9s 33ms/step - loss: 5.1888
- accuracy: 0.1438
Epoch 66/150
266/266 [==============================] - 9s 33ms/step - loss: 5.1656
- accuracy: 0.1444
Epoch 67/150
266/266 [==============================] - 9s 33ms/step - loss: 5.1585
- accuracy: 0.1477
Epoch 68/150
266/266 [==============================] - 9s 33ms/step - loss: 5.1350
- accuracy: 0.1474
Epoch 69/150
266/266 [==============================] - 9s 33ms/step - loss: 5.1179
- accuracy: 0.1489
Epoch 70/150
266/266 [==============================] - 9s 33ms/step - loss: 5.0978
- accuracy: 0.1490
Epoch 71/150
266/266 [==============================] - 9s 33ms/step - loss: 5.0742
- accuracy: 0.1482
Epoch 72/150
266/266 [==============================] - 9s 33ms/step - loss: 5.0575
- accuracy: 0.1542
Epoch 73/150
266/266 [==============================] - 9s 33ms/step - loss: 5.0374
- accuracy: 0.1549
Epoch 74/150
266/266 [==============================] - 9s 33ms/step - loss: 5.0199
- accuracy: 0.1568
Epoch 75/150
266/266 [==============================] - 9s 33ms/step - loss: 5.0103
- accuracy: 0.1554
```

```
Epoch 76/150
266/266 [==============================] - 9s 33ms/step - loss: 4.9892
- accuracy: 0.1578
Epoch 77/150
266/266 [==============================] - 9s 33ms/step - loss: 4.9721
- accuracy: 0.1582
Epoch 78/150
266/266 [==============================] - 9s 33ms/step - loss: 4.9461
- accuracy: 0.1616
Epoch 79/150
266/266 [==============================] - 9s 33ms/step - loss: 4.9283
- accuracy: 0.1597
Epoch 80/150
266/266 [==============================] - 9s 32ms/step - loss: 4.9202
- accuracy: 0.1614
Epoch 81/150
266/266 [==============================] - 9s 32ms/step - loss: 4.9028
- accuracy: 0.1633
Epoch 82/150
266/266 [==============================] - 9s 33ms/step - loss: 4.8773
- accuracy: 0.1667
Epoch 83/150
266/266 [==============================] - 9s 33ms/step - loss: 4.8687
- accuracy: 0.1665
Epoch 84/150
266/266 [==============================] - 9s 33ms/step - loss: 4.8439
- accuracy: 0.1662
Epoch 85/150
266/266 [==============================] - 9s 33ms/step - loss: 4.8385
- accuracy: 0.1673
Epoch 86/150
266/266 [==============================] - 9s 33ms/step - loss: 4.8102
- accuracy: 0.1707
Epoch 87/150
266/266 [==============================] - 9s 33ms/step - loss: 4.8025
- accuracy: 0.1690
Epoch 88/150
266/266 [==============================] - 9s 33ms/step - loss: 4.7793
- accuracy: 0.1712
Epoch 89/150
266/266 [==============================] - 9s 33ms/step - loss: 4.7625
- accuracy: 0.1699
Epoch 90/150
266/266 [==============================] - 9s 35ms/step - loss: 4.7494
- accuracy: 0.1700
Epoch 91/150
266/266 [==============================] - 9s 33ms/step - loss: 4.7410
- accuracy: 0.1728
Epoch 92/150
```

```
266/266 [==============================] - 9s 33ms/step - loss: 4.7317
- accuracy: 0.1745
Epoch 93/150
266/266 [==============================] - 9s 33ms/step - loss: 4.7049
- accuracy: 0.1708
Epoch 94/150
266/266 [==============================] - 9s 33ms/step - loss: 4.6993
- accuracy: 0.1759
Epoch 95/150
266/266 [==============================] - 9s 33ms/step - loss: 4.6777
- accuracy: 0.1758
Epoch 96/150
266/266 [==============================] - 9s 33ms/step - loss: 4.6559
- accuracy: 0.1766
Epoch 97/150
266/266 [==============================] - 9s 33ms/step - loss: 4.6497
- accuracy: 0.1768
Epoch 98/150
266/266 [==============================] - 9s 33ms/step - loss: 4.6391
- accuracy: 0.1768
Epoch 99/150
266/266 [==============================] - 9s 33ms/step - loss: 4.6255
- accuracy: 0.1781
Epoch 100/150
266/266 [==============================] - 9s 33ms/step - loss: 4.6055
- accuracy: 0.1824
Epoch 101/150
266/266 [==============================] - 9s 33ms/step - loss: 4.5904
- accuracy: 0.1804
Epoch 102/150
266/266 [==============================] - 9s 33ms/step - loss: 4.5795
- accuracy: 0.1807
Epoch 103/150
 15/266 [>.............................] - ETA: 8s - loss: 4.4103 -
accuracy: 0.1854
```

```
model.save('nwp.h5')

vocab_array = np.array(list(tokenizer.word_index.keys()))

vocab_array
```

Making Predictions

This function takes an input text and predicts the next n_words in the sequence using the trained model. It tokenizes the input text, pads it to match the input shape of the model, makes a prediction, converts the prediction index to its corresponding word, and appends the predicted word to the input text. Finally, it returns the predicted text.

```
def make_prediction(text, n_words):
    for i in range(n_words):
        text_tokenize = tokenizer.texts_to_sequences([text])
        text_padded =
tf.keras.preprocessing.sequence.pad_sequences(text_tokenize,
maxlen=14)
        prediction = np.squeeze(np.argmax(model.predict(text_padded),
axis=-1))
        prediction = str(vocab_array[prediction - 1])
        text += " " + prediction
    return text

make_prediction("cloudy", 5)
```

REFERENCES

- https://www.analyticsvidhya.com/blog/2023/07/next-word-prediction-with-bidirectional-lstm/
- https://www.youtube.com/watch?v=-roEcxiRXOA&ab_channel=PritishMishra
- https://www.youtube.com/watch?v=VB7bbFIEAhk&t=163s&ab_channel=IGTechTeam