

Codsoft Internship Aug 2023

Task -Movie Gener classification

Author- Vivek Kumar

```
In [8]: import numpy as np
import pandas as pd
import sklearn
import tensorflow as tf
```

```
In [10]: df = pd.read_csv("train.csv")
df
```

```
Out[10]:
```

	id	movie_name	synopsis	genre
0	44978	Super Me	A young scriptwriter starts bringing valuable ...	fantasy
1	50185	Entity Project	A director and her friends renting a haunted h...	horror
2	34131	Behavioral Family Therapy for Serious Psychiat...	This is an educational video for families and ...	family
3	78522	Blood Glacier	Scientists working in the Austrian Alps discov...	scifi
4	2206	Apat na anino	Buy Day - Four Men Widely - Apart in Life - By...	action
...
53995	80325	Untitled Monkeyshine/Ugly Brother Project	Two loser pest control buddies find themselves...	scifi
53996	84712	The Great Swindle	A seductive woman gets involved in relationshi...	thriller
53997	69801	When Men Get Pregnant	Duyen, a wedding dress staff, who decides to c...	romance
53998	7678	Suttu Pidikka Utharavu	The people of a crowded colony in Coimbatore c...	action
53999	33970	Margo the Mouse	Margo is a little mouse that lives quietly in ...	family

54000 rows × 4 columns

```
In [11]: df['genre'].value_counts()
```

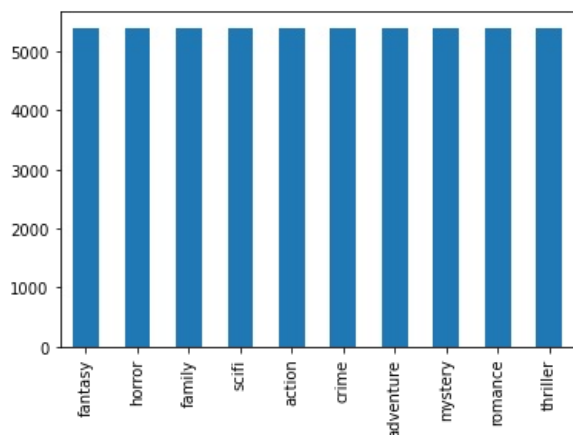
```
Out[11]:
```

fantasy	5400
horror	5400
family	5400
scifi	5400
action	5400
crime	5400
adventure	5400
mystery	5400
romance	5400
thriller	5400

Name: genre, dtype: int64

```
In [12]: (df['genre'].value_counts()).plot(kind='bar')
```

```
Out[12]: <AxesSubplot:>
```



```
In [13]: from sklearn.model_selection import train_test_split

(train_df, test_df) = train_test_split(df, test_size=0.1,
                                       random_state=42,
                                       stratify=df['genre'])

(train_df.shape, test_df.shape)
```

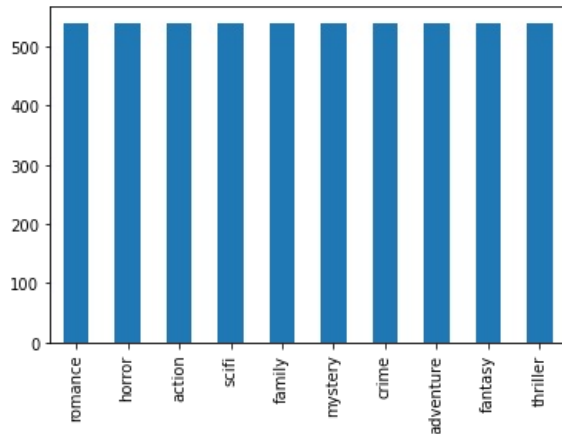
```
Out[13]: ((48600, 4), (5400, 4))
```

```
In [14]: test_df['genre'].value_counts()
```

```
Out[14]: romance      540
horror      540
action      540
scifi       540
family      540
mystery     540
crime       540
adventure   540
fantasy     540
thriller    540
Name: genre, dtype: int64
```

```
In [15]: (test_df['genre'].value_counts()).plot(kind='bar')
```

```
Out[15]: <AxesSubplot:>
```



```
In [16]: train_df, val_df = train_test_split(train_df, test_size=0.2,
                                             random_state=42,
                                             stratify=train_df['genre'])

len(train_df), len(val_df)
```

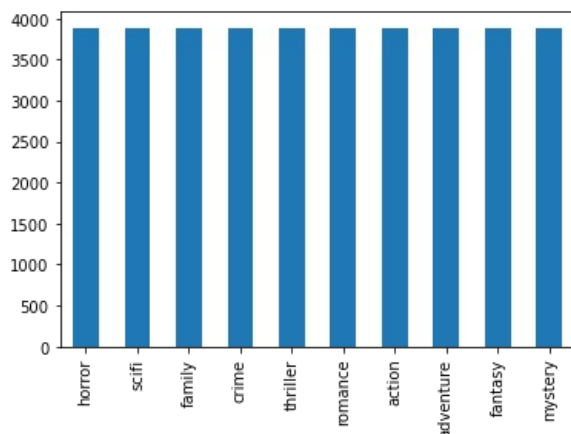
```
Out[16]: (38880, 9720)
```

```
In [17]: train_df['genre'].value_counts()
```

```
Out[17]: horror      3888
scifi      3888
family     3888
crime      3888
thriller   3888
romance    3888
action     3888
adventure  3888
fantasy    3888
mystery    3888
Name: genre, dtype: int64
```

```
In [18]: train_df['genre'].value_counts().plot(kind='bar')
```

```
Out[18]: <AxesSubplot:>
```



```
In [19]: val_df['genre'].value_counts()
```

```
Out[19]: romance      972
         crime       972
         scifi       972
         fantasy     972
         family      972
         mystery     972
         adventure   972
         action      972
         thriller    972
         horror      972
         Name: genre, dtype: int64
```

```
In [20]: movie_name_lengths = [len(name.split()) for name in train_df['movie_name']]
         movie_name_lengths = pd.DataFrame(movie_name_lengths)

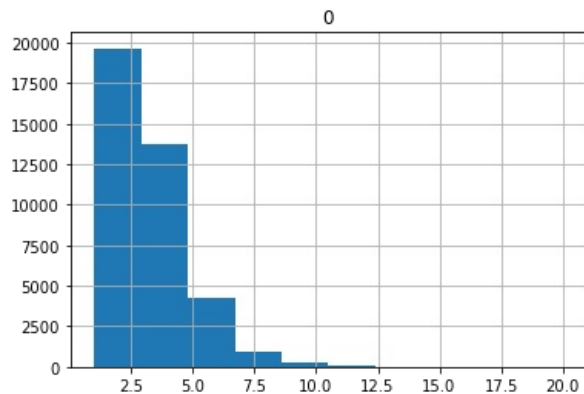
         movie_name_lengths.describe().transpose()
```

```
Out[20]:
```

	count	mean	std	min	25%	50%	75%	max
0	38880.0	2.837912	1.676202	1.0	2.0	2.0	4.0	20.0

```
In [21]: movie_name_lengths.hist()
```

```
Out[21]: array([[<AxesSubplot:title={'center':'0'}>]], dtype=object)
```



```
In [22]: np.percentile(movie_name_lengths, 99)
```

```
Out[22]: 8.0
```

```
In [23]: synopsis_lengths = [len(s.split()) for s in train_df['synopsis']]
         synopsis_lengths = pd.DataFrame(synopsis_lengths)

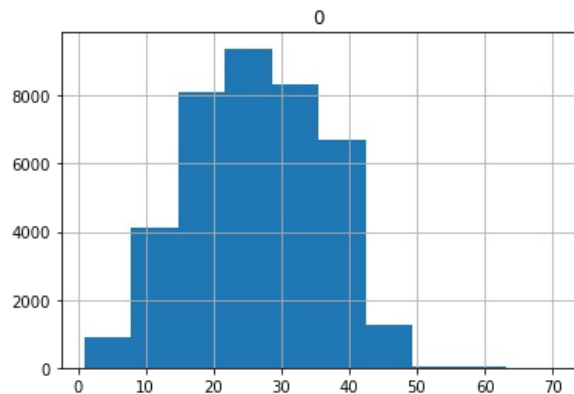
         synopsis_lengths.describe().transpose()
```

```
Out[23]:
```

	count	mean	std	min	25%	50%	75%	max
0	38880.0	26.233693	9.809758	1.0	19.0	26.0	34.0	70.0

```
In [24]: synopsis_lengths.hist()
```

```
Out[24]: array([[<AxesSubplot:title={'center':'0'}>]], dtype=object)
```



```
In [25]: from sklearn.preprocessing import LabelEncoder
```

```
         class_encoder = LabelEncoder()
         class_encoder.fit(train_df['genre'])
         CLASSES = class_encoder.classes_
         CLASSES
```

```
Out[25]: array(['action', 'adventure', 'crime', 'family', 'fantasy', 'horror',
               'mystery', 'romance', 'scifi', 'thriller'], dtype=object)
```

```
In [26]: class_encoder.transform(['action', 'crime', 'thriller'])
```

```
Out[26]: array([0, 2, 9])
```

```
In [27]: from tensorflow.data import Dataset, AUTOTUNE
```

```
BATCH_SIZE = 128
def build_model_datasets(features, labels):
    y = class_encoder.transform(labels)
    dataset = Dataset.from_tensor_slices((features, y)).batch(BATCH_SIZE).prefetch(AUTOTUNE)

    return dataset

train_dataset = build_model_datasets(train_df['synopsis'], train_df['genre'])
val_dataset = build_model_datasets(val_df['synopsis'], val_df['genre'])

train_dataset, val_dataset, len(train_dataset), len(val_dataset)
```

```
Out[27]: (<_PrefetchDataset element_spec=(TensorSpec(shape=(None,), dtype=tf.string, name=None), TensorSpec(shape=(None,
), dtype=tf.int32, name=None))>,
<_PrefetchDataset element_spec=(TensorSpec(shape=(None,), dtype=tf.string, name=None), TensorSpec(shape=(None,
), dtype=tf.int32, name=None))>,
304,
76)
```

```
In [28]: for i in train_dataset:
        print(i)
        break;
```

```
(<tf.Tensor: shape=(128,), dtype=string, numpy=
array([b'A maniac cut and burn the body of the victims without a real reason. Annie (Frances Williams) the sist
er of one of the girls killed, wanna know more about this mistery and for her starts a nightmare.',
      b'A young undercover cop fights to control a constantly changing reality while trying to maintain her ow
n grip on sanity when she joins the hunt for a brutal serial killer.',
      b'In the not too distant future, Earth is attacked by a foreign enemy. With help from an alien ally, man
kind must rebuild on a new planet but one man assigned to the job has a personal agenda that could jeopardize t
he entire mission.',
      b'It's a journey of a man who was born handicapped without hands but he never lost confidence in himself
and ends up an opened school on his mother name where all normal and abnormal kids can study together.",
      b'Val Barber, a private investigator, is hired by a wealthy widow to find her missing granddaughter.',
      b'When a bunch of kids on the naughty list receive nothing for Christmas, their parents decide to sue Sa
nta Claus.',
      b'A former circus artist escapes from a mental hospital to rejoin his armless mother - the leader of a s
trange religious cult - and is forced to enact brutal murders in her name as he becomes "her arms".',
      b'A paranoid man begins videotaping the strange, unexplainable events happening around him.',
      b'As Ishaan, a music producer who has commitment issues with the women in his life, feels a connection w
ith Gauri, his next-door neighbor, he begins to learn the importance of relationships.',
      b'The jury in a New York City murder trial is frustrated by a single member whose skeptical caution forc
es them to more carefully consider the evidence before jumping to a hasty verdict.',
      b'Max Kirkpatrick is a cop who protects Kate McQuean, a civil law attorney, from a renegade KGB team out
to terminate her.',
      b'In 1968 five young sleuths must decode a series of complex clues based on Thomas Edison's life and inv
entions in order to find a local tycoon's hidden wealth, hoping the money will save their coal mining town from
impending financial ruin.",
      b'The young people who met years later at the University and fell in love with each other are actually t
he children of two hostile families, but they will not give up their love.',
      b'A New York fashion model finds herself being pursued by a poor but honest garage mechanic and a rich p
hilanderer.',
      b'Santhana Devan, a freedom fighter, secretly operates a gang to act against the British government. His
cover is blown when his own sister is raped and murdered by a British officer.',
      b'A group of friends find their party is being crashed by a group of mysterious cat-like ghouls. To make
matters worse, their school's murderous bullies decide to show up as well.",
      b'The Supreme Court orders the release of three Satanist men convicted of raping Jake Cabrera's wife Han
nah. Jake decides to render his own brand of justice. However, the souls of the three convicts possess him, cau
sing harm to his family.",
      b'Charlie and Lisa, two divorced parents in their 40's who find themselves at a midlife crossroads. Both
are single parents and they have four complicated teenagers.",
      b'A mad female scientist needs Santo's blood for a youth serum; and she has Frankenstein's reliable old
monster to help her get it.",
      b'Black Easter is based on the famous legend of a bride that was killed on her wedding night. The ghostl
y bride also known as "Resurrection Mary" hitchhikes and kills everyone in her path on each anniversary of that
fateful night.',
      b'In a gothic setting, a detective uses a mysterious machine to solve one of the Century's most difficul
t cases.",
      b'Plot has not been disclosed at this time.',
      b'When a long lost tribe stakes its claim to Animal Paradise, Yugo, Lala and Uncle Bear team up once mor
e, this time with parental supervision.',
      b'In this origin story of Father Christmas, an ordinary boy (with a loyal pet mouse and a reindeer at hi
s side) sets out on an extraordinary adventure to find his father who is on a quest to discover the fabled vill
age of Elfhelm.',
      b'A group of college students decide to shoot a documentary on the psychological causes that lead to a f
orm of demonic possession. However, the boys will remain trapped in their game of suggestion with tragically un
expected effects.',
      b'The businessman who sends the first Chilean astronaut into outer space can't quite come up with the mo
ney to bring the adventurer back.",
      b'Two powerful female warriors must journey across a post-apocalyptic land to fight a ritual duel, and f
ulfill an ancient prophecy.',
      b'Three legendary, immortal warriors reinvent themselves as modern-day crime-fighters.',
      b'A computer genius is confined to a penthouse after a car accident, with only a nurse to look after him
```

. But the advanced computer systems designed to make his life easier appear to be taking control.',
 b'A witch put to death in 1692 swears vengeance on her persecutors and returns to the present day to punish their descendants.',
 b'An urban husband and wife travel to the jungle and learn just how precious their relationship is.',
 b'US sleuth Whispering Smith arrives in London on vacation, but soon becomes embroiled in a murder case.',
 ,
 b"Allied agents stage a daring raid on a castle where the Nazis are holding American brigadier general George Carnaby prisoner, but that's not all that's really going on.",
 b'They know when you are alone',
 b'Young Jesse travels through time trying to stop a bunch of evil, virtual, fight managers from destroying Earth.',
 b"Bamse and his friends go to the Arctic to rescue Skelman's scientist friend Beanka who has gone missing. On the way they are joined by the deceitful Kattja who wants the reward money.",
 b'After a wealthy man finds out his wife is a gold digger, he meets and falls for a widowed mother, and complications ensue.',
 b'Alisa Freyndlikh and Andrey Podoshian star in this enchanting retelling of Hans Christian Andersen's classic fairy tale. When a handsome prince tries to find his princess bride, he searches high and low, only to discover love at his doorstep.',
 b'A warrior returning home to his country must battle giant bats, three-headed dogs and a vicious dragon to save his wife, and his people, from the machinations of an evil ruler.',
 b'A love story between a "demi-mondaine" kept by various lovers, suffering from tuberculosis, and a young provincial bourgeois. Adaptation from the novel by Alexandre Dumas, fils, "The Lady of the Camellias".',
 b'Advait visits Goa where he meets Sara, a free-spirited girl who lives life unshackled. Opposites attract and all goes well until life turns upside down. Years later, Advait is on a killing spree with cops Aghase and Michael in his way.',
 b'A stand up comedian with a brutal alter-ego is interrogated by police about the contents of his hard drive.',
 b'Two FBI agents attempt to clarify the murders occurring in a desolate region. They approach the witnesses of the latest incident with the help of the local police. All of them hide something and all have wildly different stories to tell.',
 b'After the tragic loss of their child, a couple must prepare to witness their last day of light before the world fades into ever winter and darkness.',
 b'Anand comes to Chennai with an ambition of becoming an actor. However, his life turns upside down when he gets embroiled in the murder of a minister.',
 b'Arthur C Clarke wrote: "Mankind will build a tower from the earth's surface into space about 50 years after everyone has stopped laughing about it." That time is here.',
 b'A truly moving and cruel story, which denounces some of the many injustices that occur in Venezuelan prisons, where the law of the strongest prevails.',
 b'The Christmas-themed family film will follow Pat Keeley, a young man from Ireland who has emigrated to the United States and finds himself caught up in an escalating whirlwind of shenanigans that takes him from New York to Santa Fe.',
 b'A young man joins a gang of smugglers that looks for someone with clean criminal record.',
 b'The story of two men with very different lives but united for the pleasure of killing.',
 b'In a french lost village an old woman is killed and her savings stolen. Several members of her family, all called:"Goupi-...", are suspected.',
 b'A rich kid, with a fondness for fast bikes, is blamed when a motorcycle accident puts his best friend in a coma.',
 b'In 2072 Rome, Italy, the ruthless CEO of a TV network plots to stage a modern series of gladiator games for ratings while one 'contentant' learns of a conspiracy behind the true nature of the results.',
 b'A young woman's life turns chaotic when she uncovers a deadly secret about her friend.',
 b'Police officer Yukimura Sayo and her younger sister Miwa visit a ryokan located deep in the mountains. They don't know that this traditional inn famous for a nearby hot spring with healing qualities is a den for vampires.',
 b'After a peace-loving Japanese immigrant and his family become victims of a crime syndicate, a master ninja emerges.',
 b'On his way to become the world's richest, a young and corrupted billionaire named Alain Arnault is hunted by an anthropomorphic being known as "The God of Money".',
 b'Vacationing Englishman Oliver Larker stumbles upon a plot to kill a young American tourist in Spain.',
 b'Ho is a refugee from Hong Kong. He strives with talent and hardworking, and grows into a casino tycoon.',
 ,
 b"A widowed reporter recruits the help of a federal agent to investigate her late husband's secrets, but the two become the target of unknown attackers.",
 b'Corporate Shadow Ops Commander Cassandra Lake learns that her young daughter holds a secret that her employer must protect at all costs. Sandra must protect her daughter from the overwhelming forces rallied against her.',
 b"Strange things happen when Chandru and Sangeetha decide to get married. To his horror, Chandru learns that they are being pursued and attacked by the ghost of Sangeetha's mother from her previous life.",
 b'A retired cop and his new partner must put their differences aside while investigating a drug ring.',
 b'A father travels from Oklahoma to France to help his estranged daughter, who is in prison for a murder she claims she didn't commit.",
 b'While doing research in the woods, four college friends find a cursed Indian burial mound and find themselves trapped in a mysterious maze, followed by a terrifying force.',
 b'A nerd spends his time at an allotment shed away from his wife. When the zombie apocalypse begins, he and his nerdy friend try to make it through and save the wife and her cute friend.',
 b'In the near future, convicts are offered the chance to volunteer as medical subjects to shorten their sentence. One such subject for a new drug capable of generating feelings of love begins questioning the reality of his emotions.',
 b'Joan flees New York City for the sweltering canyons above Los Angeles after her married lover's suicide and she plots to take down a woman from her haunted past.",
 b'With the help of an inventor's secret workshop, Ritter Rost tries to lead his country out of a terrible crisis and become a knight again.",
 b'A group of British and Aboriginal slaves in 1800's Australia, work together to find an ancient spear and take back their land.",
 b'A woman goes undercover as an exotic dancer to kill the mob boss that murdered her boyfriend.',
 b'A Bounty Hunter and his private network of elite vigilantes must prevent a dangerous crime syndicate from establishing absolute power.',
 b'Recently orphaned, a young boy is taken in by his godmother who is shocked to realize that she can see the boy's imaginary friend: a flamboyant, French magician named Bogus.",
 b'A woman named Yeon-hee (Ha Ji-won) lives in Busan with her boyfriend Man-sik (Sol Kyung-gu) near Haeun

dae Beach. But, when they find out a tsunami will hit the city, They realize they only have 10 minutes to escape!'

b'Five friends go onto a streamyads live stream panel like they usually do, however, there were meant to be Six of them, turns out one of them is dead and the killer is on the panel. Then things start to get terrifying.'

b'An American woman inherits an abandoned apartment in Paris and searches to uncover the mystery of why her Great Aunt kept it secret for over seventy years.'

b'It's Pinnocchio's sixth birthday, and Geppetto, the Good Fairy, and Pinnocchio's friends throw him a magical surprise party. Includes two puppet stories - "Puss in Boots" and "The Drummer Boy", X',

b'Dad is left home with the kids when their Mom is away for the weekend, but when Dad has to work overtime at the Chocolate Factory, mayhem ensues when the kids get into a feud with the nasty boys next door!'

b'In the film, St. Nicolas' Eve threatens to be canceled because a young Piet drops St. Nicholas's clock. In addition, the Saint's horse is behaving very strangely because he has run into Britt Dekker.",

b'In the aftermath of a mysterious and violent farm incident, a public health student seeks to uncover its origins, but her efforts run into direct conflict with the interests of those in power.'

b'A famous novelist breaks out of a mental hospital, taking an ex pro quarterback who fell from grace ho stage. The two men embark on a crazy road trip of self discovery through the Arizona desert.'

b'Based on the prologue from the popular novel, People of Paper follows the exploits of a talented former monk who had been famed for his abilities in origami.'

b'During WWII, the British Royal Navy uses a captured German submarine to trick the enemy and protect their supply convoy.'

b'Two sisters inherit their family castle, which is said to be haunted by their ancestor, a dark-haired, red-robed woman who is said to take seven lives every hundred years.'

b'There is a haunted house in Nandigama. Dead bodies of 34 people resurface in house's well over a period of time. A TV channel plans to host a reality show in that haunted house.",

b'A young widow defies society's norms to marry the man of her choosing.",

b'Once again the legendary tomb explorer Hu Bayi is on a dangerous mission as he seeks out the Tomb of Emperor Xian, located on an island of monstrous creatures in this mystical action-adventure.'

b'Alexandre Dupr  is a con artist, accidentally involved in a hunt for microfilm, containing a valuable invention of a recently murdered French scientist.'

b'Kermit and his friends go to New York City to get their musical on Broadway only to find it's a more difficult task than they anticipated.",

b'After discovering that the reason behind all of his physical differences is a genetic mutation called "Marfan Syndrome", Mo is forced to come to terms with a more serious reality.'

b'The parents of a girl who was killed by a savage dog are granted the opportunity to spend three days with their deceased daughter.'

b'A soldier wakes up in someone else's body and discovers he's part of an experimental government program to find the bomber of a commuter train within 8 minutes.",

b'A woman offers her country home to friends to help them save their marriage, but her well-meaning gesture creates a new problem.'

b'A dreamer unhappy with his job makes an impulsive decision to become a commercial trucker and tries convincing his brother to join him.'

b'A high school student falls in love with a Nordic tourist.'

b'A geologist aids the escape of a young woman whose father was murdered to recover forged oil documents.'

b'This film is called Fantasia 3 because it contains three fairy tales. The second of the three is The Wonderful Wizard of Oz.'

b'After seeing his ex-girlfriend (Henderson) turn down a nationally televised marriage proposal, a small-time crook (Carlyle) returns to his hometown to try and win back her heart.'

b'Plot outline not being disclosed at this time.',

b'Follows a young Viking as he aspires to hunt dragons, and how he becomes unexpectedly a friend of a young dragon.'

b'American family inherits a ski resort in Lapland and it's not the only thing they inherit.",

b'A flight team is captured and held captive. They are run through a gauntlet of tests to see what they will give up and what they won't.",

b'In the year 2044, a mutation of the plague brought annihilation to three-quarters of humanity. Those who remain have organized themselves in distant shelters, outside, gangs of looters control the territory.'

b'A cyborg policewoman and her partner hunt a mysterious and powerful hacker called the Puppet Master.'

b'An atheist park ranger and a bible salesman find themselves thrown into a slimy conspiracy that threatens to destroy not just the woods - but the entire world.'

b'A CIA Agent who is unsure whom he can trust, must find and destroy a weaponized nerve agent - that targets specific races and ethnicity - before it is developed and used in a deadly international conspiracy for world domination.'

b'Money's needed for Kjeld's silver wedding and Egon Olsen has an ingenious plan - steal a red bag with DKK5000000 from the crooked VP of Denmark's largest insurance company. Will Egon survive it?",

b'The adventures of an investigator (Cagney) for the Bureau of Weights and Measures.'

b'Follows a group of five teens heading off to a cabin for a "romantic" weekend. They are hunted down by The Lashman for being unruly teens and disturbing his peace.'

b'An insomnia plagued loner becomes increasingly obsessed with an elusive rat that he believes has invaded his home.'

b'A neo-noir revenge thriller centering on Julia Shames, who after suffering a brutal trauma, falls prey to an unorthodox form of therapy to restore herself.'

b'A group of college students break into an abandoned convent and become possessed by demonic spirits.'

b'A British expedition formed by the lead researcher Jonathan Marchant, his assistant, a doctor and a TV crew, travels to Congo to seek evidence of a dinosaur.'

b'Diehard Packers fan, Ben 'Zagger' Zagowski, buys a house in the parking lot of Lambeau Field and is forced to choose between a superfan lifestyle and his fianc , Amy. Based on a true story, set during the 2009 football season.",

b'Kees is an 11-year-old boy who lives in the center of the city of Amsterdam in the early 20th century.'

b'Young Dr. Kildare tries to help an unlucky brain surgeon and his seemingly insane patient.'

b'Two unlikely companions face gun-slinging cowboys and their own personal family conflicts as they learn the importance of friendship, loyalty, and forgiveness.'

b'A university research scientist, about to lose funding and status, has a lab accident and discovers he can see people's true intentions -- making his situation even worse.",

b'Daisuke struggles with a traumatic accident. One day, he met Ayumi who's being harassed by yakuza. He decides to protect her from them but he developed of seeing violence as a way to gain the affirmation and approval of others,"

b'Misaki is a schoolgirl who must fight gangs.'

b'A small-town newspaperman begins to suspect that a wave of murders committed in the area may not be the work of a serial killer but a monster.',
 b'Comedy centering on the question of whether a man's wife is or isn't pregnant.",
 b'In 1880, a former Gunfighter, is released from Prison and sets out to reunite with a lost love, but finds himself befriended by five strangers as his journey leads him into a deadly confrontation with an old nemesis.',
 b'Young Catherine Morelli (Julie Christie), who lives in Rome, goes to Geneva to find romance at her father's wedding. There she begins a nearly nymphomaniacal pursuit of a mystery-fantasy man called Gregory (Michael Sarrazin).",
 b'Five years ago Frank Phoenix robbed The Firm. Betrayed by his own crew and murdered, the stolen gold was never found. Now someone is planning revenge.',
 b'We invite you to see "UP ON THE ROOF" an extremely engaging, gripping love story with a dose of comedy, drama and numerous twists and turns.',
 b'The crew of a space ship confronts an evil galactic ruler out to rule the universe.',
 b'A worldly ambitious Monsignor clashes with his older brother, a cynical Los Angeles homicide detective who is investigating the brutal murder of a young prostitute.'],
 dtype=object)>, <tf.Tensor: shape=(128,), dtype=int32, numpy=
 array([5, 5, 8, 3, 2, 3, 9, 5, 7, 2, 7, 3, 7, 7, 0, 5, 5, 7, 1, 5, 8, 5,
 3, 3, 5, 8, 1, 0, 2, 5, 4, 6, 1, 6, 8, 1, 7, 4, 0, 4, 2, 5, 2, 6,
 0, 8, 2, 3, 0, 9, 6, 0, 8, 9, 0, 0, 4, 6, 9, 9, 8, 5, 9, 2, 5, 5,
 9, 9, 3, 1, 0, 0, 3, 9, 6, 6, 4, 3, 3, 8, 1, 8, 1, 6, 5, 7, 4, 0,
 7, 4, 5, 8, 7, 3, 7, 2, 4, 7, 7, 1, 8, 8, 8, 6, 0, 0, 2, 6, 9, 5,
 5, 5, 0, 3, 3, 2, 1, 6, 0, 0, 2, 7, 3, 6, 0, 7, 1, 2])>)

```
In [29]: from tensorflow.keras.layers import TextVectorization
```

```
SYNOPSIS_VOCAB_SIZE = 4096
SYNOPSIS_SEQ_LENGTH = 64
```

```
vectorizer = TextVectorization(max_tokens=SYNOPSIS_VOCAB_SIZE, output_sequence_length=SYNOPSIS_SEQ_LENGTH)
vectorizer.adapt(train_df['synopsis'])
```

```
In [30]: vectorizer.get_vocabulary()[:5], vectorizer.get_vocabulary()[-5:]
```

```
Out[30]: (['', '[UNK]', 'a', 'the', 'to'],
          ['freespirted', 'fred', 'filipino', 'fed', 'fair'])
```

```
In [31]: vectorizer(train_df['synopsis'][0])
```

```
Out[31]: <tf.Tensor: shape=(64,), dtype=int64, numpy=
array([ 2, 23, 1, 298, 1806, 2083, 1, 70, 19, 8, 648,
       1105, 5, 90, 1947, 15, 2, 561, 2405, 50, 317, 32,
       327, 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, 0, 0, 0, 0,
       0, 0, 0, 0, 0, 0, 0, 0, 0], dtype=int64)>
```

```
In [32]: from tensorflow.keras import Sequential
         from tensorflow.keras.layers import Embedding, Dense, Input, GlobalAveragePooling1D
```

```
EMBEDDING_LENGTH = 256
def build_model_0():
    return Sequential([
        Input(shape=(1,), dtype=tf.string),
        vectorizer,
        Embedding(SYNOPSIS_VOCAB_SIZE, EMBEDDING_LENGTH, input_length=SYNOPSIS_SEQ_LENGTH),
        GlobalAveragePooling1D(),
        Dense(256, activation='relu'),
        Dense(len(CLASSES), activation='softmax')
    ], name='model_0')
```

```
In [33]: from tensorflow.keras.losses import SparseCategoricalCrossentropy
         from tensorflow.keras.optimizers import Adam
```

```
def compile_model(model):
    model.compile(
        loss=SparseCategoricalCrossentropy(),
        optimizer=Adam(),
        metrics=['accuracy']
    )
```

```
In [34]: model_0 = build_model_0()
         compile_model(model_0)
         model_0.summary()
```

Model: "model_0"

Layer (type)	Output Shape	Param #
text_vectorization (TextVectorization)	(None, 64)	0
embedding (Embedding)	(None, 64, 256)	1048576
global_average_pooling1d (GlobalAveragePooling1D)	(None, 256)	0
dense (Dense)	(None, 256)	65792
dense_1 (Dense)	(None, 10)	2570

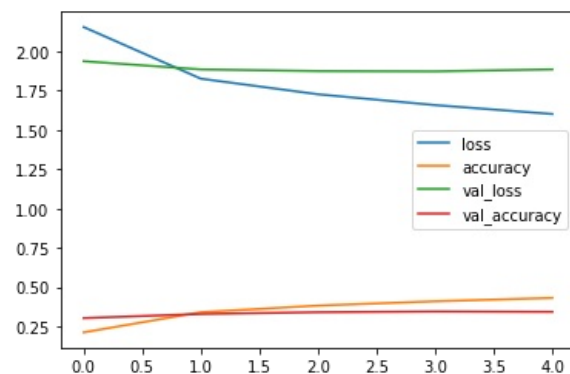
=====
Total params: 1116938 (4.26 MB)
Trainable params: 1116938 (4.26 MB)
Non-trainable params: 0 (0.00 Byte)
=====

```
In [35]: model_0_history = model_0.fit(train_dataset, epochs=5, validation_data=val_dataset)
```

```
Epoch 1/5
304/304 [=====] - 6s 17ms/step - loss: 2.1525 - accuracy: 0.2129 - val_loss: 1.9356 - val_accuracy: 0.3035
Epoch 2/5
304/304 [=====] - 5s 17ms/step - loss: 1.8247 - accuracy: 0.3408 - val_loss: 1.8836 - val_accuracy: 0.3296
Epoch 3/5
304/304 [=====] - 5s 17ms/step - loss: 1.7256 - accuracy: 0.3829 - val_loss: 1.8729 - val_accuracy: 0.3408
Epoch 4/5
304/304 [=====] - 5s 17ms/step - loss: 1.6567 - accuracy: 0.4102 - val_loss: 1.8711 - val_accuracy: 0.3455
Epoch 5/5
304/304 [=====] - 5s 17ms/step - loss: 1.6003 - accuracy: 0.4316 - val_loss: 1.8837 - val_accuracy: 0.3433
```

```
In [36]: pd.DataFrame(model_0_history.history).plot()
```

```
Out[36]: <AxesSubplot:>
```



```
In [37]: X_TEST = Dataset.from_tensor_slices(test_df['synopsis']).batch(BATCH_SIZE).prefetch(AUTOTUNE)
X_TEST
```

```
Out[37]: <_PrefetchDataset element_spec=TensorSpec(shape=(None,), dtype=tf.string, name=None)>
```

```
In [38]: Y_TEST = class_encoder.transform(test_df['genre'])
batched_Y_TEST = Dataset.from_tensor_slices(Y_TEST).batch(BATCH_SIZE).prefetch(AUTOTUNE)
batched_Y_TEST
```

```
Out[38]: <_PrefetchDataset element_spec=TensorSpec(shape=(None,), dtype=tf.int32, name=None)>
```

```
In [42]: from sklearn.metrics import precision_recall_fscore_support
from sklearn.metrics import ConfusionMatrixDisplay

def model_predict(model):
    predictions = model.predict(X_TEST)
    predictions = np.argmax(predictions, axis=1)
    predictions = class_encoder.inverse_transform(predictions)
    return predictions

def plot_confusion_matrix(y_true, y_pred):
    display = ConfusionMatrixDisplay.from_predictions(y_true, y_pred,
                                                       cmap='Greens')

    fig = display.ax_.get_figure()
    fig.set_figwidth(8)
    fig.set_figheight(8)

def evaluate_model(model):
```

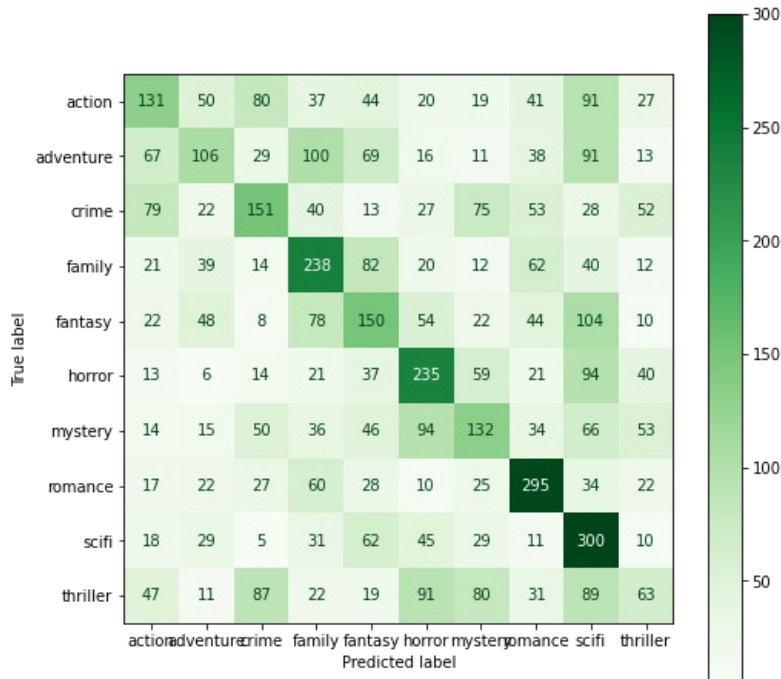


```

predictions = model_predict(model)
macro_metrics = precision_recall_fscore_support(test_df['genre'], predictions, average="macro")
weighted_metrics = precision_recall_fscore_support(test_df['genre'], predictions, labels=CLASSES)
plot_confusion_matrix(test_df['genre'], predictions)
return {
    'weighted_metrics': weighted_metrics,
    'precision': macro_metrics[0],
    'recall': macro_metrics[1],
    'f1-score': macro_metrics[2],
    'predictions': predictions
}
model_0_data = evaluate_model(model_0)

```

43/43 [=====] - 0s 4ms/step



In [43]: model_0_data

```

Out[43]: {'weighted_metrics': (array([0.30536131, 0.3045977 , 0.32473118, 0.35897436, 0.27272727,
    0.38398693, 0.28448276, 0.46825397, 0.32017076, 0.20860927]),
    array([0.24259259, 0.1962963 , 0.27962963, 0.44074074, 0.27777778,
    0.43518519, 0.24444444, 0.5462963 , 0.55555556, 0.11666667]),
    array([0.27038184, 0.23873874, 0.30049751, 0.39567747, 0.27522936,
    0.40798611, 0.26294821, 0.5042735 , 0.40622884, 0.14964371]),
    array([540, 540, 540, 540, 540, 540, 540, 540, 540, 540, 540], dtype=int64)),
    'precision': 0.3231895505247933,
    'recall': 0.3335185185185185,
    'f1-score': 0.3211605289171138,
    'predictions': array(['romance', 'horror', 'action', ..., 'fantasy', 'horror', 'action'],
    dtype=object)}

```

```

In [44]: def report_model(data):
    print("===== Metrics =====")
    print(f"{'class':16s} |{'precision':>11}|{'recall':>11}|{'F1-score':>11}|")
    for i, label in enumerate(CLASSES):
        metrics = data['weighted_metrics']
        precision = metrics[0][i]
        recall = metrics[1][i]
        f1 = metrics[2][i]
        print(f"{'label.upper():16s} |{precision:>11.6f}|{recall:>11.6f}|{f1:>11.6f}|")

    print("=====")
    print(f"{'GLOBAL':16s} |{data['precision']:>11.6f}|{data['recall']:>11.6f}|{data['f1-score']:>11.6f}|")
    print("=====")

report_model(model_0_data)

```

Metrics			
class	precision	recall	F1-score
ACTION	0.305361	0.242593	0.270382
ADVENTURE	0.304598	0.196296	0.238739
CRIME	0.324731	0.279630	0.300498
FAMILY	0.358974	0.440741	0.395677
FANTASY	0.272727	0.277778	0.275229
HORROR	0.383987	0.435185	0.407986
MYSTERY	0.284483	0.244444	0.262948
ROMANCE	0.468254	0.546296	0.504274
SCIFI	0.320171	0.555556	0.406229
THRILLER	0.208609	0.116667	0.149644
GLOBAL	0.323190	0.333519	0.321161

In []:

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js