

# Лабораторна робота №4

Хід роботи:

## Завдання на лабораторну роботу 4

1. Використати будь-який датасет з реченнями, які можна поділити на класи: позитивний текст, негативний текст, нейтральний текст.
2. Підготувати датасет до аналізу: токенизувати (перетворити слова у цифри) речення за допомогою `Tokenizer()` та створити послідовності (перетворити речення у послідовність цифр, де кожна цифра відповідає певному слову) за допомогою `tokenizer.text_to_sequence()` методу. Доповнити послідовності нулями до однакової довжини за допомогою методу `pad_sequences()`.
3. Розділити дані на тренувальні та тестові.
4. Реалізувати рекурентну нейронну мережу LSTM для аналізу тону тексту. Нейронна мережа має на вхід отримувати речення, та на виході повертати тон тексту: позитивний негативний або нейтральний.
5. Натренувати нейронну мережу та протестувати, відобразити результати тестування.

Код:

```
import numpy as np
from tensorflow.keras.preprocessing.text import Tokenizer
from tensorflow.keras.preprocessing.sequence import pad_sequences
from tensorflow.keras.utils import to_categorical
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import Embedding, LSTM, Dense, SpatialDropout1D
from sklearn.model_selection import train_test_split
import matplotlib.pyplot as plt

sentences = [
    "I love this product, it is fantastic!",
    "This is the worst movie I have ever seen.",
    "The service was okay, nothing special.",
    "I am very happy with my purchase.",
    "I am disappointed with the quality.",
    "It's just an average experience.",
```

					ДУ «Житомирська політехніка».23.121.03.000 – Лр4			
Змн.	Арк.	№ докум.	Підпис	Дата	Звіт з лабораторної роботи	Лім.	Арк.	Аркушів
Розроб.	Бігун Р.В.						1	7
Перевір.	Годлевський Ю.О.					ФІКТ Гр. ВТ-22-1[1]		
Керівник								
Н. контр.								
Зав. каф.								

```

"Amazing performance by the lead actor!",
"I wouldn't recommend this to anyone.",
"The food was delicious and the service was great.",
"This product broke after one use.",
"It's neither good nor bad, just alright.",
"The book was a pleasant surprise, I enjoyed it a lot.",
"The customer service was terrible.",
"The event was well-organized and fun.",
"I don't think I'll be coming back here again.",
"It was an okay movie, not the best but not the worst either.",
"Absolutely fantastic! Exceeded my expectations.",
"Very poor quality, I'm really disappointed.",
"The hotel was clean and comfortable.",
"The wait time was too long.",
"The presentation was neither exciting nor boring.",
"Great value for the price, highly recommend!",
"I'm upset with the service I received.",
"The software is user-friendly and intuitive.",
"The app crashes frequently, not reliable at all.",
"It's a pretty average tool, does what it needs to.",
"Outstanding job by the team, well done!",
"The taste was awful, couldn't finish my meal.",
"The lecture was informative and engaging.",
"I'm never buying from this store again.",
"The class was okay, not too difficult but not too easy.",
"Excellent quality and fast shipping!",
"The package arrived damaged and late.",
"It's a decent place, nothing extraordinary.",
"Loved the atmosphere and the friendly staff.",
"The product did not meet my expectations at all.",
"The trip was fun, though the weather was bad.",
"The design is sleek and modern, very satisfied.",
"The fabric feels cheap and uncomfortable.",
"The course content was satisfactory, covered the basics.",
"Impressed with the customer support!",
"Terrible experience, I want a refund.",
"The meeting was somewhat productive.",
"Best purchase I've made this year.",
"The software has too many bugs.",
"The restaurant ambiance was nice, food was okay.",
"Wonderful product, will buy again.",
"The installation process was confusing and hard.",
"The show was decent, had some good moments.",
]

labels = [1, 0, 2, 1, 0, 2, 1, 0, 1, 0, 2, 1, 0, 1, 0, 2, 1, 0, 1, 0, 2, 1, 0,
1, 0, 2, 1, 0, 1, 0, 2, 1, 0, 2, 1, 0, 2, 1, 0, 2, 1, 0, 2, 1, 0, 2]

max_words = 10000

```

					ДУ «Житомирська політехніка».23.121.03.000 – Пр4	Арк.
						2
Змн.	Арк.	№ докум.	Підпис	Дата		

```

tokenizer = Tokenizer(num_words=max_words)
tokenizer.fit_on_texts(sentences)
sequences = tokenizer.texts_to_sequences(sentences)

# Доповнення послідовностей нулями до однакової довжини
maxlen = 20
data = pad_sequences(sequences, maxlen=maxlen)

num_classes = 3
labels = to_categorical(labels, num_classes)

X_train, X_test, y_train, y_test = train_test_split(data, labels,
test_size=0.2, random_state=42)

# Побудова LSTM моделі
model = Sequential()
model.add(Embedding(max_words, 128, input_length=maxlen))
model.add(SpatialDropout1D(0.2))
model.add(LSTM(100, dropout=0.2, recurrent_dropout=0.2))
model.add(Dense(num_classes, activation='softmax'))

model.compile(loss='categorical_crossentropy', optimizer='adam',
metrics=['accuracy'])

batch_size = 32
epochs = 30

history = model.fit(X_train, y_train, batch_size=batch_size, epochs=epochs,
validation_data=(X_test, y_test), verbose=1)

score, acc = model.evaluate(X_test, y_test, batch_size=batch_size, verbose=2)
print('Test score:', score)
print('Test accuracy:', acc)

# Відображення результатів тестування
plt.plot(history.history['accuracy'], label='train')
plt.plot(history.history['val_accuracy'], label='test')
plt.title('Model accuracy')
plt.ylabel('Accuracy')
plt.xlabel('Epoch')
plt.legend(['Train', 'Test'], loc='upper left')
plt.show()

plt.plot(history.history['loss'], label='train')
plt.plot(history.history['val_loss'], label='test')
plt.title('Model loss')
plt.ylabel('Loss')

```

					ДУ «Житомирська політехніка».23.121.03.000 – Лр4	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		3

```

plt.xlabel('Epoch')
plt.legend(['Train', 'Test'], loc='upper left')
plt.show()

def predict_sentiment(text, model, tokenizer, maxlen):
    sequences = tokenizer.texts_to_sequences([text])
    padded_sequences = pad_sequences(sequences, maxlen=maxlen)
    predictions = model.predict(padded_sequences)
    # Перетворення ймовірностей у відсотки
    percentages = predictions[0] * 100
    return percentages

# Приклад використання функції
text = "This is amazing!"
percentages = predict_sentiment(text, model, tokenizer, maxlen)
class_labels = ['Negative', 'Positive', 'Neutral']

for i, label in enumerate(class_labels):
    print(f"{label}: {percentages[i]:.2f}%")

```

**Вигляд:**

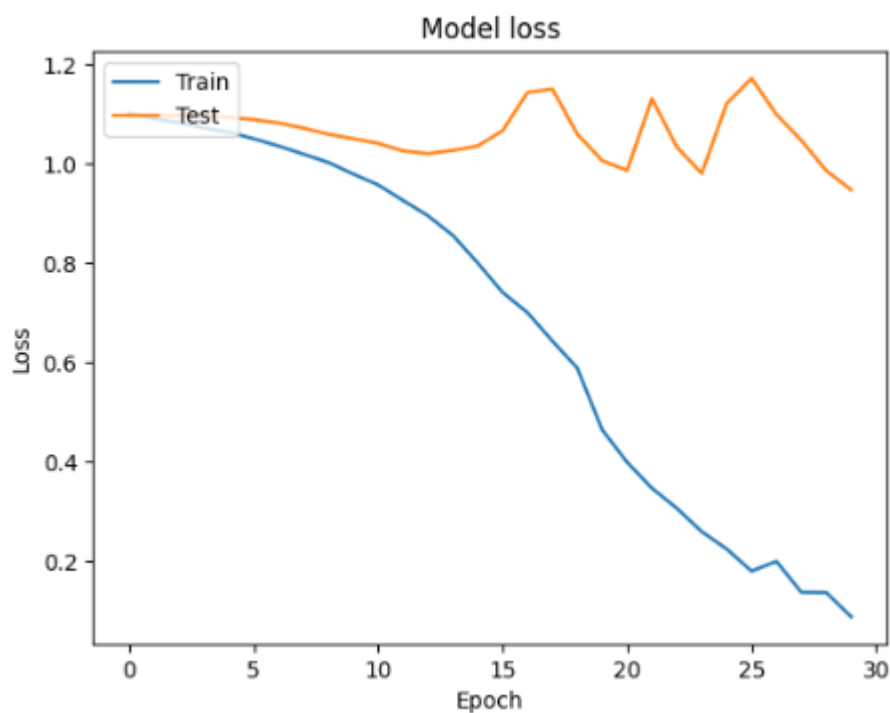
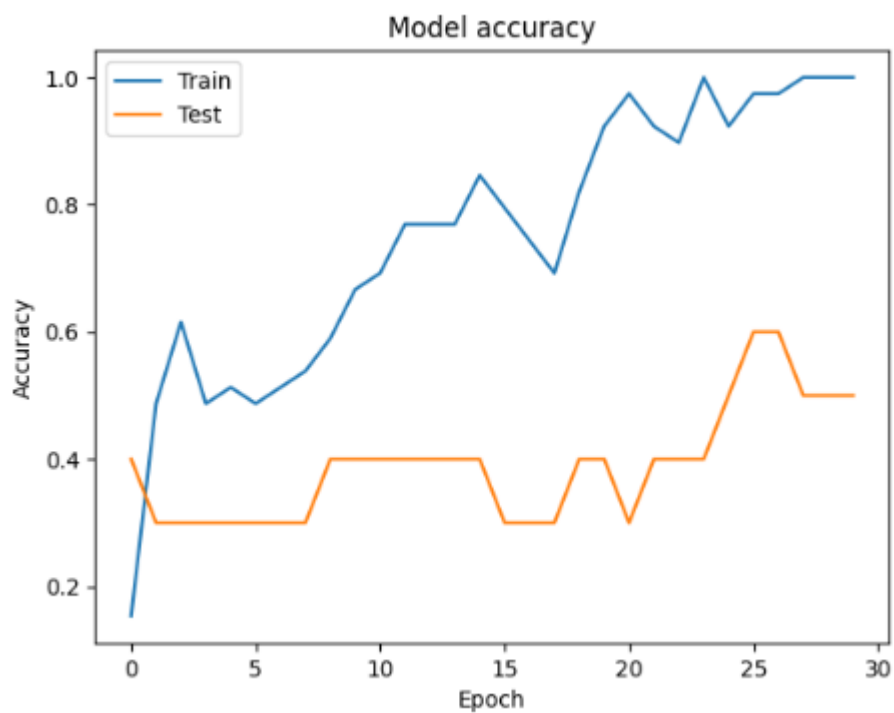
					ДУ «Житомирська політехніка».23.121.03.000 – Лр4	Арк.
Змн.	Арк.	№ докум.	Підпис	Дата		4

```

2/2 [=====] - 0s 87ms/step - loss: 1.0731 - accuracy: 0.4872 - val_loss: 1.0976 - val_accuracy: 0.3000
Epoch 5/30
2/2 [=====] - 0s 99ms/step - loss: 1.0642 - accuracy: 0.5128 - val_loss: 1.0942 - val_accuracy: 0.3000
Epoch 6/30
2/2 [=====] - 0s 78ms/step - loss: 1.0511 - accuracy: 0.4872 - val_loss: 1.0895 - val_accuracy: 0.3000
Epoch 7/30
2/2 [=====] - 0s 75ms/step - loss: 1.0364 - accuracy: 0.5128 - val_loss: 1.0829 - val_accuracy: 0.3000
Epoch 8/30
2/2 [=====] - 0s 77ms/step - loss: 1.0199 - accuracy: 0.5385 - val_loss: 1.0724 - val_accuracy: 0.3000
Epoch 9/30
2/2 [=====] - 0s 78ms/step - loss: 1.0033 - accuracy: 0.5897 - val_loss: 1.0601 - val_accuracy: 0.4000
Epoch 10/30
2/2 [=====] - 0s 80ms/step - loss: 0.9799 - accuracy: 0.6667 - val_loss: 1.0511 - val_accuracy: 0.4000
Epoch 11/30
2/2 [=====] - 0s 74ms/step - loss: 0.9580 - accuracy: 0.6923 - val_loss: 1.0417 - val_accuracy: 0.4000
Epoch 12/30
2/2 [=====] - 0s 75ms/step - loss: 0.9267 - accuracy: 0.7692 - val_loss: 1.0266 - val_accuracy: 0.4000
Epoch 13/30
2/2 [=====] - 0s 74ms/step - loss: 0.8959 - accuracy: 0.7692 - val_loss: 1.0210 - val_accuracy: 0.4000
Epoch 14/30
2/2 [=====] - 0s 100ms/step - loss: 0.8563 - accuracy: 0.7692 - val_loss: 1.0275 - val_accuracy: 0.4000
Epoch 15/30
2/2 [=====] - 0s 77ms/step - loss: 0.8007 - accuracy: 0.8462 - val_loss: 1.0365 - val_accuracy: 0.4000
Epoch 16/30
2/2 [=====] - 0s 79ms/step - loss: 0.7413 - accuracy: 0.7949 - val_loss: 1.0673 - val_accuracy: 0.3000
Epoch 17/30
2/2 [=====] - 0s 89ms/step - loss: 0.7005 - accuracy: 0.7436 - val_loss: 1.1443 - val_accuracy: 0.3000
Epoch 18/30
2/2 [=====] - 0s 79ms/step - loss: 0.6434 - accuracy: 0.6923 - val_loss: 1.1512 - val_accuracy: 0.3000
Epoch 19/30
2/2 [=====] - 0s 75ms/step - loss: 0.5887 - accuracy: 0.8205 - val_loss: 1.0603 - val_accuracy: 0.4000
Epoch 20/30
2/2 [=====] - 0s 75ms/step - loss: 0.4644 - accuracy: 0.9231 - val_loss: 1.0070 - val_accuracy: 0.4000
Epoch 21/30
2/2 [=====] - 0s 87ms/step - loss: 0.3991 - accuracy: 0.9744 - val_loss: 0.9871 - val_accuracy: 0.3000
Epoch 22/30
2/2 [=====] - 0s 88ms/step - loss: 0.3467 - accuracy: 0.9231 - val_loss: 1.1317 - val_accuracy: 0.4000
Epoch 23/30
2/2 [=====] - 0s 76ms/step - loss: 0.3061 - accuracy: 0.8974 - val_loss: 1.0342 - val_accuracy: 0.4000
Epoch 24/30
2/2 [=====] - 0s 82ms/step - loss: 0.2590 - accuracy: 1.0000 - val_loss: 0.9812 - val_accuracy: 0.4000
Epoch 25/30
2/2 [=====] - 0s 74ms/step - loss: 0.2242 - accuracy: 0.9231 - val_loss: 1.1217 - val_accuracy: 0.5000
Epoch 26/30
2/2 [=====] - 0s 74ms/step - loss: 0.1796 - accuracy: 0.9744 - val_loss: 1.1727 - val_accuracy: 0.6000
Epoch 27/30
2/2 [=====] - 0s 93ms/step - loss: 0.1993 - accuracy: 0.9744 - val_loss: 1.1001 - val_accuracy: 0.6000
Epoch 28/30
2/2 [=====] - 0s 94ms/step - loss: 0.1369 - accuracy: 1.0000 - val_loss: 1.0478 - val_accuracy: 0.5000
Epoch 29/30
2/2 [=====] - 0s 100ms/step - loss: 0.1363 - accuracy: 1.0000 - val_loss: 0.9868 - val_accuracy: 0.5000
Epoch 30/30
2/2 [=====] - 0s 90ms/step - loss: 0.0877 - accuracy: 1.0000 - val_loss: 0.9483 - val_accuracy: 0.5000
1/1 - 0s - loss: 0.9483 - accuracy: 0.5000 - 28ms/epoch - 28ms/step
Test score: 0.948264479637146
Test accuracy: 0.5

```

					ДУ «Житомирська політехніка».23.121.03.000 – Пр4	Арк.
						5
Змн.	Арк.	№ докум.	Підпис	Дата		



```
# Приклад використання функції
text = "This is amazing!"
percentages = predict_sentiment(text, model, tokenizer, maxlen)
class_labels = ['Negative', 'Positive', 'Neutral']

for i, label in enumerate(class_labels):
    print(f"{label}: {percentages[i]:.2f}%")
```

```
1/1 [=====] - 0s 276ms/step
Negative: 18.74%
Positive: 80.17%
Neutral: 1.09%
```

**GIT HUB:** [https://github.com/ShadowGhost31/Lab\\_4](https://github.com/ShadowGhost31/Lab_4)

**Висновки:** в ході виконання лабораторної роботи я реалізував нейронну мережу LSTM для аналізу тону тексту. Протестував її та відобразив і дослідив результати тестування.

					ДУ «Житомирська політехніка».23.121.03.000 – Лр4	Арк.
						7
Змн.	Арк.	№ докум.	Підпис	Дата		