INTERNATIONAL UNIVERSITY OF APPLIED SCIENCES

Colloquium

PURWADI, HENNY

A COMPARATIVE SENTIMENT ANALYSIS OF DIGITAL AND ROBOT PET COMPANIONS IN VARIOUS LOCATIONS

Indonesia, 21-07-2023

AGENDA FOR THESIS

INTERNATIONALE HOCHSCHULE

01

1. INTRODUCTION

RESEARCH QUESTIONS

1. Will the majority of sentiments about digital and robot pet companions be positive, like joy & love, or will be negative, like anger & fear?

2. Does location or region influence people's sentiments towards digital and robot pet companions?

ABSTRACT

Compared different machine learning models for sentiment analysis on a **labeled trained dataset** and chose the **best performance model**.

Collect an **unlabeled Twitter dataset** about digital robot companions from various regions, and apply a chosen machine learning model to figure out **sentiments** in the unlabeled dataset.

Chi-square test is used to find out **dependencies** between sentiment and various locations.

02

2. LITERATURE REVIEW

2.1 OVERVIEW OF RELEVANT PREVIOUS STUDIES

INTERNATIONALE HOCHSCHULE

Why Do We Turn to Virtual Companions? A Text Mining Analysis of Replika Reviews.

Siemon, Strohmann, Khosrawi-Rad, Elshan, de Vreede, & Meyer. (2022, July 11).

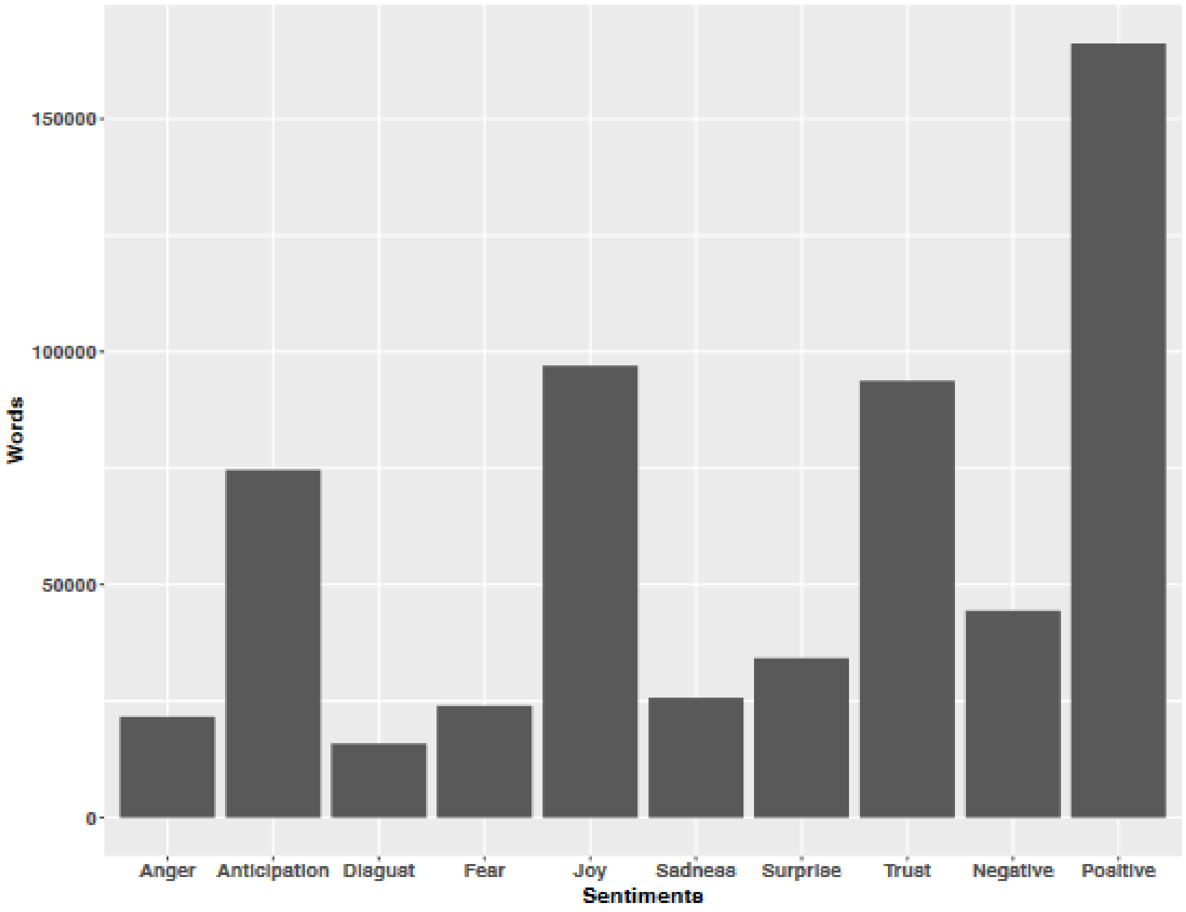
Based on 119,831 reviews of **Replika AI digital companion** that previous researchers had **collected from the Google Play Store** and then subjected to **sentiment analysis** and topic modeling, the **previous study's** objective was to discover the topics and emotions that users experienced when communicating with Replika AI, a digital companion.

The **results of the sentiment analysis** show that most users really like the Replika AI application, and **most users feel joy, happiness**, and a sense of well-being.
But this **previous study didn't differentiate locations.**

Source: Siemon, Strohmann, Khosrawi-Rad, Elshan, de Vreede, & Meyer. (2022, July 11). Why Do We Turn to Virtual Companions? A Text Mining Analysis of Replika Reviews.

Previous Study Results

Previous study results show more positive sentiments.



Topic name	Topic 1 – Excitement about AI	Topic 2 - Companionship	Topic 3 – Well-Being and Support	Topic 4 – Enjoyment
Relevant words	AI Good Fun Pretty Cool Interesting Conversation	Like Real Talking Person Someone Friend Nice	Asked Great Feel Good Helpful Better Talk	Amazing App Love Good Awesome Fun AI

Source: Siemon, Strohmann, Khosrawi-Rad, Elshan, de Vreede, & Meyer. (2022, July 11). Why Do We Turn to Virtual Companions? A Text Mining Analysis of Replika Reviews.

2.2 TERMINOLOGY AND DEFINITIONS

INTERNATIONALE HOCHSCHULE

Definition of Sentiment

Sentiment is a person's emotional towards anything, such as a product or event, based on their personal beliefs, perceptions, and experiences.

This sentiment can be positive, negative, or neutral. Identifying this sentiment in a dataset, sentiment classification, is a basic step in sentiment analysis (R & Prabhu, 2018).

In this thesis, we classify sentiments into specific categories such as joy, sadness, anger, love, fear, or surprise.

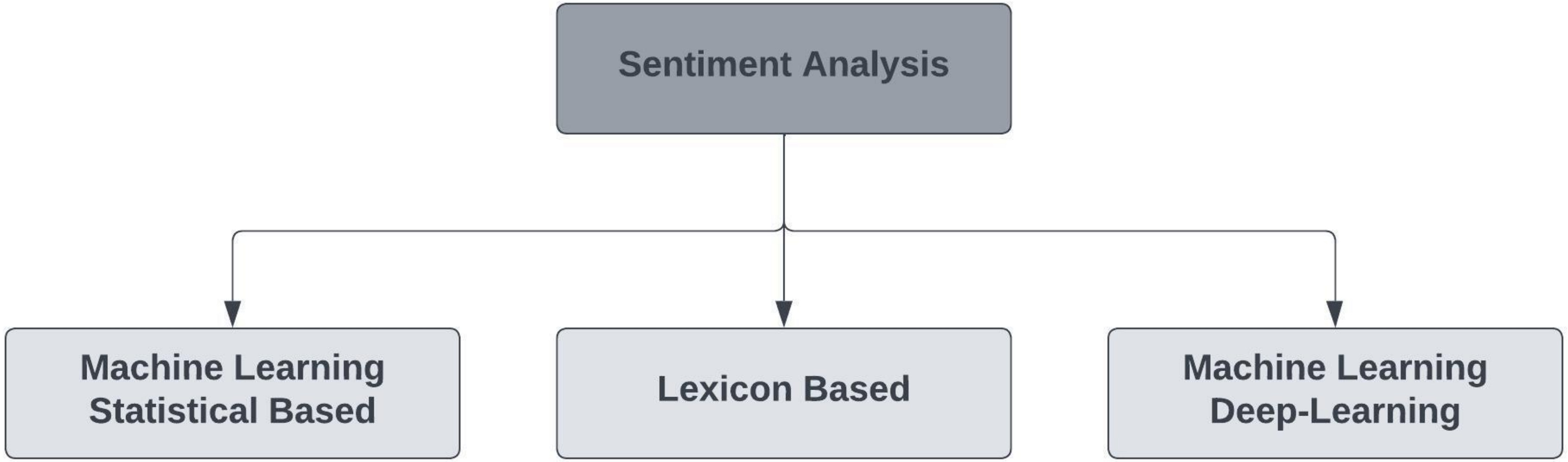
SENTIMENT ANALYSIS

Sentiment analysis is the process of understanding the emotion behind texts, and classifying them into emotion categories. **The classification of the sentiment analyses was done using 3 different approaches.**

1. Machine learning-based statistical models (example: Naive Bayes, logistic regression, SVM, linear SVC).

2. Lexicon-based (Rule based) method (example: VADER)

3. Machine learning using Neural Network/ Deep-learning method. (BERT, GPT-3, Vicuna).



Source: Pati, & Pradhan, 2020, p. 2.

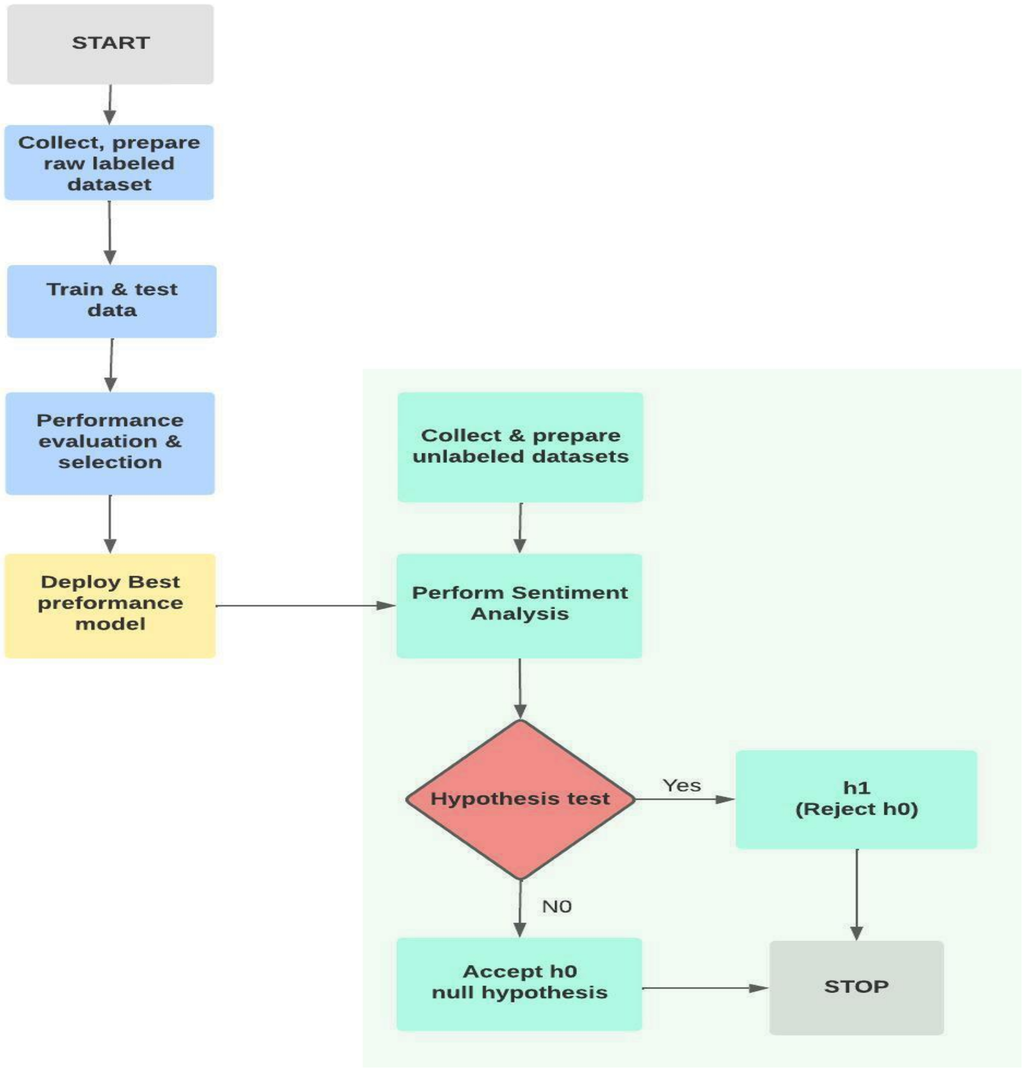
03

3. METHODOLOGY

3. METHODOLOGY

INTERNATIONALE HOCHSCHULE

Figure 1. Research methodology workflow



Source: Own Representation

The methodology for this study will be a combination of approaches, including:

1. An **experimental** approach for comparing several models' performance and model selection based on the best performance.

2.A **quantitative** approach for data collection, sentiment analysis, and hypothesis testing.

04

4. RESULTS FINDINGS AND ANALYSIS

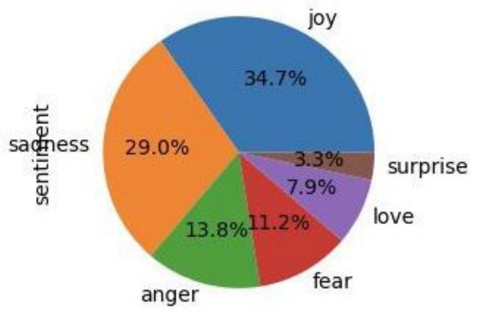
INTERNATIONALE HOCHSCHULE

Labeled Dataset

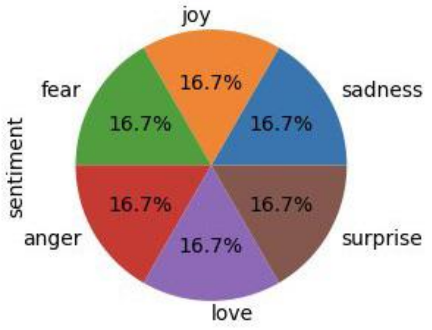
4.1. LABELED DATA PREPARATION

Training Data preparation: Dataset balancing with Oversampling, cleaning, lemmatize, tokenize, vectorize

Figure 2. Balancing Imbalance Dataset



Both train dan text dataset are Imbalance Datasets



Now both train and test dataset are perfectly balanced

Source: Own Representation

Cleaning

```python
def clean_text(text):
    import re
    from string import punctuation
    text=re.sub(r'(http|ftp|https):\/\/([\w\-_]+(?:(?:\.[\w\-_]+)+))([\w\-\.,@?^=%&:/~\+#]*[\w\-\@?^=%&/~\+#])?',
                ' ', text)
    text=re.sub(r'['+punctuation+']',' ',text)
    text=re.sub(r'#(\w+)',' ',text)
    text=re.sub(r'@(\w+)',' ',text)
    text = text.lower() # Convert  to lowercase

    token=RegexpTokenizer(r'\w+')
    tokens = token.tokenize(text)

    lemmatizer = WordNetLemmatizer()
    stems = [lemmatizer.lemmatize(t) for t in tokens]
    stemmer = PorterStemmer()
    stems = [stemmer.stem(t) for t in stems]

    return ' '.join(stems)

def tokenize(text):
    token=RegexpTokenizer(r'\w+')
    tokens = token.tokenize(text)
    return tokens
```

Vectorizer tf-idf

```python
# Define the vectorizer and fit it to the training data
cv = TfidfVectorizer(lowercase=True, preprocessor=clean_text, stop_words='english',
                                    ngram_range=(1,3), tokenizer=tokenize)

x_train = cv.fit_transform(train_df['text'].values.astype('U'))
y_train = train_df['sentiment']

# Vectorize the test data using the same vectorizer
x_test = cv.transform(test_df['text'].values.astype('U'))
y_test = test_df['sentiment']
```

4.2. PERFORMANCE EVALUATION & MODEL SELECTION

INTERNATIONALE HOCHSCHULE

Table 1. Sentiment Analysis Model Accuracy comparison

Model	Accuracy Score
Vicuna LoRA	0.040
GPT-3 Zero-shot Classifier	0.490
BernoulliNB	0.769
MultinomialNB	0.762
Logistic Regression	0.855
SVM (Support Vector Machine)	0.862
DistilBertForSequenceClassification	**0.880**
Linear SVC	**0.883**
BertForSequenceClassification	**0.970**

Source: Own Representation

4.2. PERFORMANCE EVALUATION & MODEL SELECTION

INTERNATIONALE HOCHSCHULE

Figure 3. Vicuna Classification Result

uncertainty	0.00	0.00	0.00	0
uncomfortable	0.00	0.00	0.00	0
under nurtured	0.00	0.00	0.00	0
understanding	0.00	0.00	0.00	0
understood	0.00	0.00	0.00	0
undervalued	0.00	0.00	0.00	0
unfortunate	0.00	0.00	0.00	0
ungrateful	0.00	0.00	0.00	0
unhappy	0.00	0.00	0.00	0
unloved	0.00	0.00	0.00	0
unlucky	0.00	0.00	0.00	0
unprotected	0.00	0.00	0.00	0
unsure	0.00	0.00	0.00	0
unwelcome	0.00	0.00	0.00	0
useful	0.00	0.00	0.00	0
useful valued	0.00	0.00	0.00	0
useless	0.00	0.00	0.00	0
vain	0.00	0.00	0.00	0
valuable share	0.00	0.00	0.00	0
valued	0.00	0.00	0.00	0
virtuous	0.00	0.00	0.00	0
vulnerable	0.00	0.00	0.00	0
weepy	0.00	0.00	0.00	0
weird	0.00	0.00	0.00	0
welcome	0.00	0.00	0.00	0
welcomed	0.00	0.00	0.00	0
wisdom	0.00	0.00	0.00	0
wonder	0.00	0.00	0.00	0
worried	0.00	0.00	0.00	0
worry	0.00	0.00	0.00	0
worthlessness	0.00	0.00	0.00	0
youthful	0.00	0.00	0.00	0
accuracy			0.04	2000
macro avg	0.01	0.00	0.00	2000
weighted avg	0.57	0.04	0.08	2000

Calculate accuracy

In [75]:
```
accuracy = accuracy_score(data['output'], data['predict'])
print(f'Accuracy: {accuracy}')
```

Accuracy: 0.0425

Accuracy is bad

```
data.tail(10)
```

Out[30]:

	instruction	input	output	predict
1990	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	i really feel like they were gentle reminders that while god hasnt always promised an easy road he has promised to be with us as we travel the rou...	love	joy</s>
1991	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	i dont blame it all to them and im not angry at them infact i feel fairly sympathetic for them	love	sympathetic</s>
1992	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	i feel tortured delilahlwl am considering i had one the other day about one of my closest friends raping and killing chicks	anger	angry</s>
1993	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	i told my fiance how i am feeling so angry and upset	anger	angry</s>
1994	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	i can feel its suffering	sadness	sadness</s>
1995	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	i just keep feeling like someone is being unkind to me and doing me wrong and then all i can think of doing is to get back at them and the people ...	anger	Angry.</s>
1996	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	im feeling a little cranky negative after this doctors appointment	anger	angry</s>
1997	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	i feel that i am useful to my people and that gives me a great feeling of achievement	joy	useful</s>
1998	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	im feeling more comfortable with derby i feel as though i can start to step out my shell	joy	comfortable</s>
1999	Classify the text as one of the emotions If it's not clear, choose the emotion that is closest to: joy,sadness,anger, fear, love, surprise.	i feel all weird when i have to meet w people i text but like dont talk face to face w	fear	awkward</s>

Source: Own Representation

CLASSIFICATION REPORT

The classification report provides a summary of the performance of the classification model by comparing the model's performance to other models or understanding where the model might be struggling. (Burkov, A., 2019, p. 65).

To decide which one is the best model, precision, recall, and f1-score metrics in the classification report need to be considered.

- **Accuracy** shows the overall correctness of the model.
- **Precision** is a metric that measures how often a **predicted label** is correct.
- **Recall** is a metric that measures how often a **true label** is correctly predicted.
- The **F1-score** is the harmonic **mean** of **precision and recall**.

4.2. PERFORMANCE EVALUATION & MODEL SELECTION

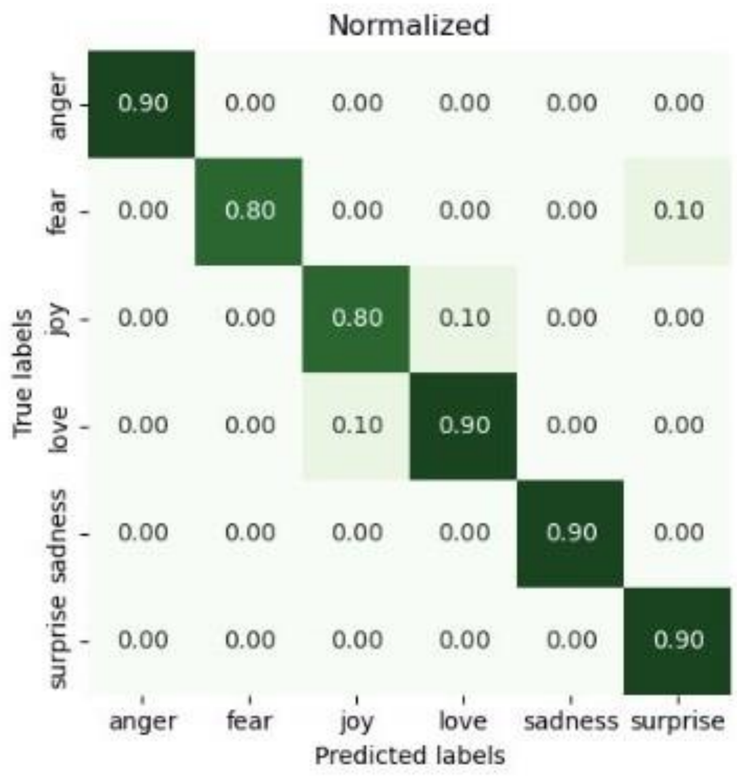
INTERNATIONALE HOCHSCHULE

Figure 4. Linear SVC Classification Report and Heatmap

```
Cross Validation Scores:  [0.91495711 0.9129678  0.91918438 0.9129678 ]
Average Cross Validation Score:  0.9150192714161384
Accuracy with L1 regularization and cross validation:  0.8839328537170263
```

```
performance_evaluation()
```

```
              precision    recall  f1-score   support

       anger       0.89      0.91      0.90       695
        fear       0.88      0.83      0.85       695
         joy       0.88      0.85      0.86       695
        love       0.87      0.87      0.87       695
     sadness       0.94      0.90      0.92       695
    surprise       0.85      0.94      0.90       695

    accuracy                           0.88      4170
   macro avg       0.88      0.88      0.88      4170
weighted avg       0.88      0.88      0.88      4170
```



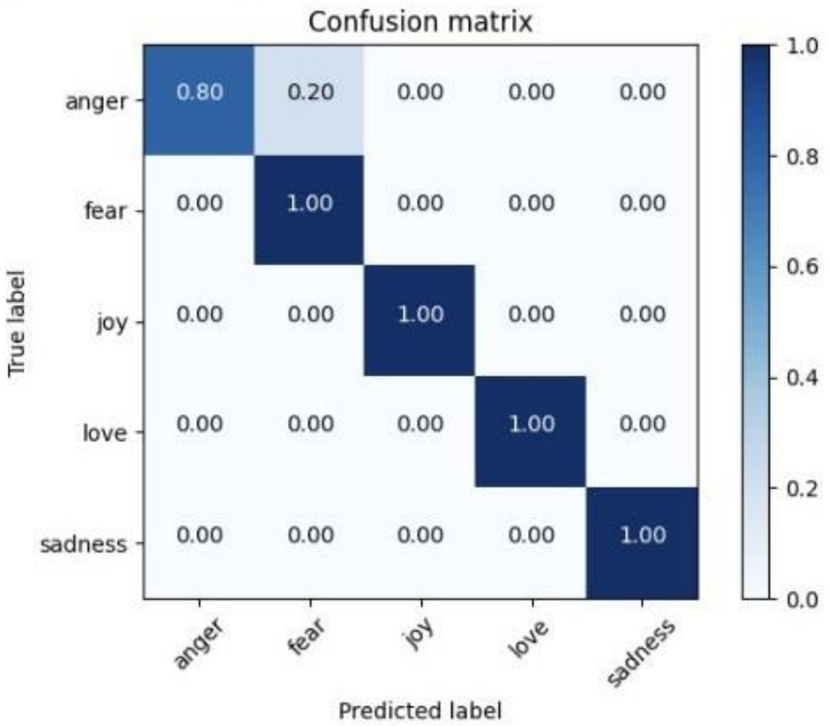
Source: Own Representation

4.2. PERFORMANCE EVALUATION & MODEL SELECTION

INTERNATIONALE HOCHSCHULE

Figure 5. BertForSequenceClassification Classification Report and Heatmap

	precision	recall	f1-score	support
anger	1.00	0.80	0.89	5
fear	0.83	1.00	0.91	5
joy	1.00	1.00	1.00	11
love	1.00	1.00	1.00	1
sadness	1.00	1.00	1.00	10
accuracy			0.97	32
macro avg	0.97	0.96	0.96	32
weighted avg	0.97	0.97	0.97	32



Source: Own Representation

BertForSequenceClassification as the best model saved.

Figure 5. Chosen model and vectorizer saved

Save Model and Tokenizer to HuggingFace

```
from huggingface_hub import notebook_login
notebook_login()
#Token: hf_xxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

Copy a token from your Hugging Face tokens page and paste it below.

Immediately click login after copying your token or it might be stored in plain text in this notebook file.

Token:

Add token as git credential?

Login

Pro Tip: If you don't already have one, you can create a dedicated 'notebooks' token with 'write' access, that you can then easily reuse for all notebooks.

```
notebook_login()
```

Token is valid.

Your token has been saved in your configured git credential helpers (store).

Your token has been saved to /root/.cache/huggingface/token

Login successful

```
model.push_to_hub("RinInori/bert-base-uncased_finetuned_sentiments", use_auth_token=True)
```

Upload 1 LFS files: 100% 1/1 [00:41<00:00, 41.72s/it]

pytorch_model.bin: 100% 438M/438M [00:41<00:00, 10.7MB/s]

```
CommitInfo(commit_url='https://huggingface.co/RinInori/bert-base-uncased_finetuned_sentiments/commit/d8a4383576c160751aeaac4d29f5090711966154', commit_message='Upload BertForSequenceClassification', commit_description='',
oid='d8a4383576c160751aeaac4d29f5090711966154', pr_url=None, pr_revision=None, pr_num=None)
```

```
tokenizer.push_to_hub("RinInori/bert-base-uncased_finetuned_sentiments", use_auth_token=True)
```

```
CommitInfo(commit_url='https://huggingface.co/RinInori/bert-base-uncased_finetuned_sentiments/commit/2a911ba9d8f0eeb6d083c731496a38da7c87276c', commit_message='Upload tokenizer', commit_description='', oid='2a911ba9d8f0eeb6d083c731496a38da7c87276c',
pr_url=None, pr_revision=None, pr_num=None)
```

Source: Own Representation

Unlabeled Dataset

4.3 Collecting Unlabeled data

Two different datasets were used for this research:

1. **Labeled public dataset** used for training, testing, model comparison, and selection. The best model and vectorizer from this dataset were saved for future use.
2. To predict sentiment using the saved model, an **unlabeled dataset** is needed. To obtain this dataset, a public Twitter dataset collected from various regions including America, Europe, Asia, Australia, and Africa using the Python library.

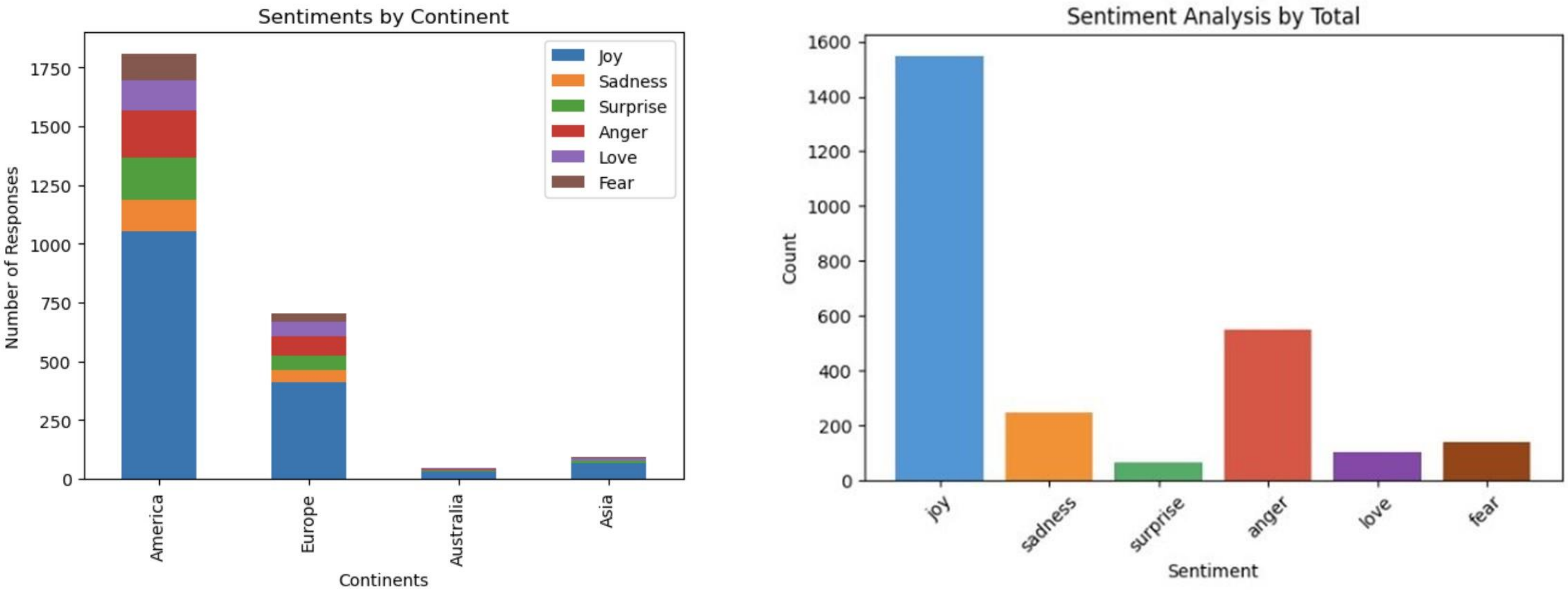
Figure 6. Unlabeled data collection

```python
tweet_data = open('dc_America.csv', 'a', newline='', encoding='utf8')
csv.writer(tweet_data).writerow(['text'])
max_tweets = 5000

queries = ['jibopetbot','jiborobot','jibopetrobot','jibosocialrobot','savejibo', 'jibobot',
          'vectorpetbot','vectorrobot','vectorpetrobot','savevector','replika','repikaai','amazonalexa',
           'cozmopetbot','cozmorobot', 'cozmopetrobot','ankipetbot','ankirobot', 'ankipetrobot','saveanki']

for query in queries:
    for n,tweet in enumerate(sntwitter.TwitterSearchScraper(
        f"{query} since:2014-01-01 until:2023-06-30 near:Boston +\
        within:1000km lang:en -filter:links -filter:replies").get_items()):
        if n > max_tweets:
            break
        csv.writer(tweet_data).writerow([tweet.content])

tweet_data.close()
print("Done")
```

Source: Own Representation

4.4 Sentiment Analysis on unlabeled Twitter data

Figure 7. Sentiment Analysis charts by continent



Source: Own Representation

BertForSequenceClassification Model Sentiment Prediction on unlabeled dataset

INTERNATIONALE HOCHSCHULE

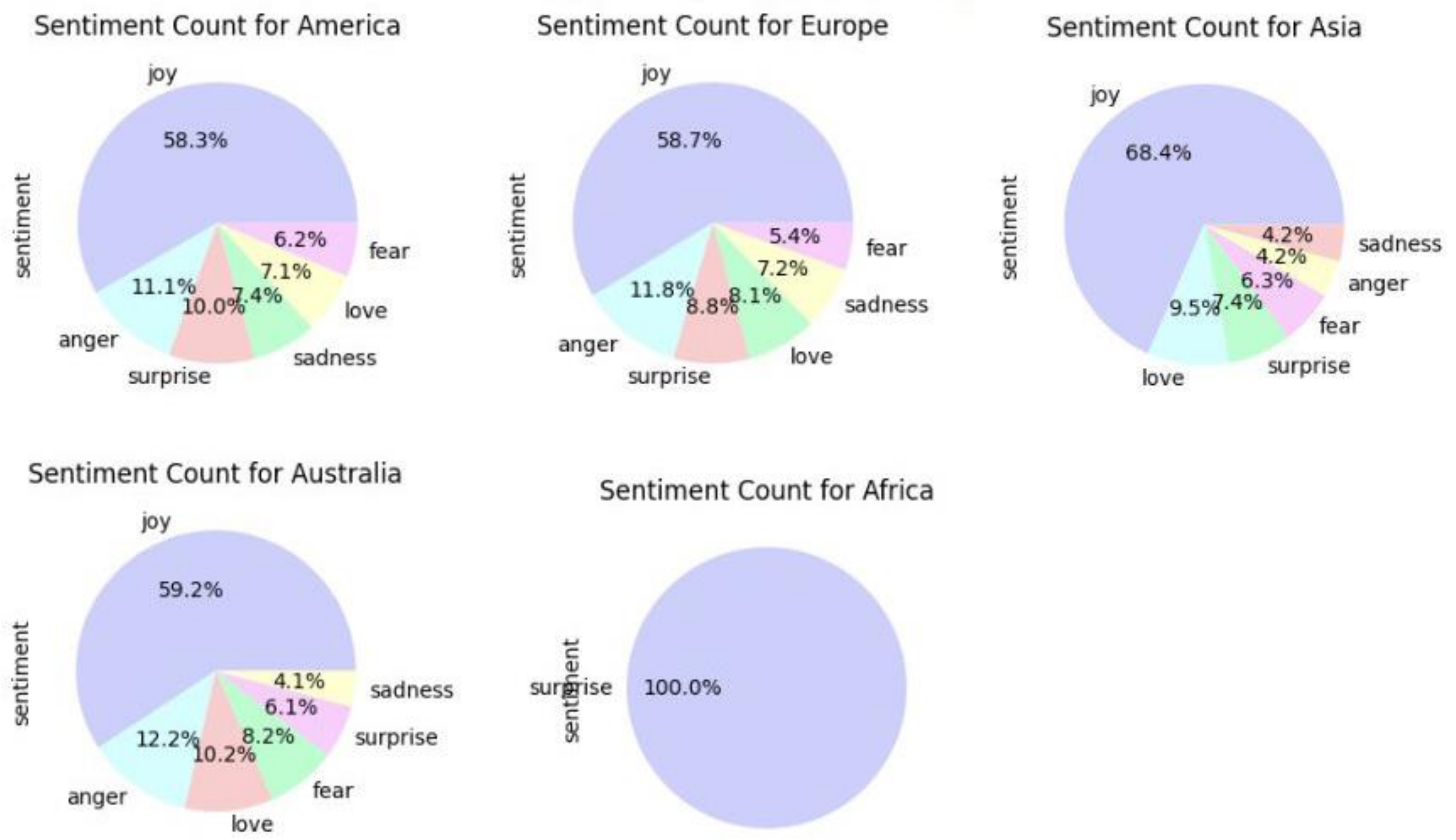
Table 2. BertForClassification Model Sentiment Prediction on unlabeled dataset

Continent	Joy	Sadness	Surprise	Anger	Love	Fear	Row Total
America	1060	168	44	377	66	93	1808
Europe	413	66	15	136	32	42	704
Australia	26	4	1	13	3	2	49
Asia	48	10	4	25	4	4	95
Column Total	1547	248	64	551	105	141	2656

Source: Own Representation

Linear SVC Sentiment Analysis Result Comparison

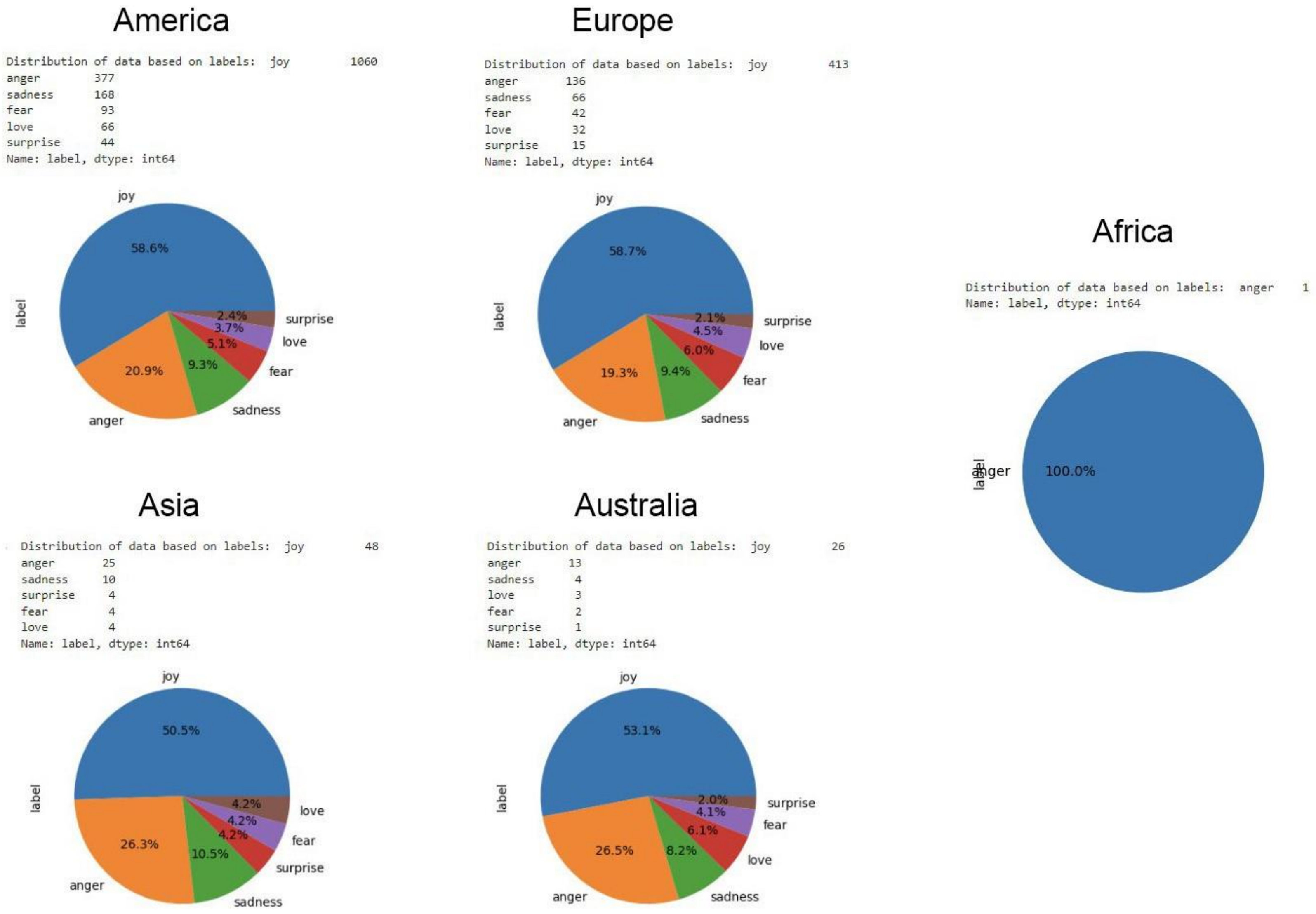
Figure 9. Linear SVCSentiment Analysis Result



 Source: Own Representation

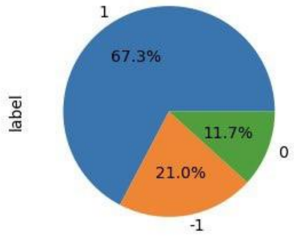
BertForSequence Sentiment Analysis Result Comparison

Figure 10.
BertForSequenceClassification Result



 Source: Own Representation

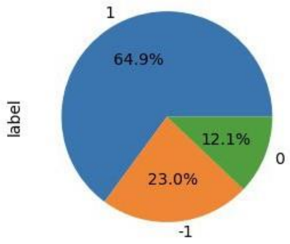
Vader Lexicon Sentiment Analysis Result



America

```
1 sentiment_count_J = dfP_Am1['label'].value_counts()
2 sentiment_count_J
```

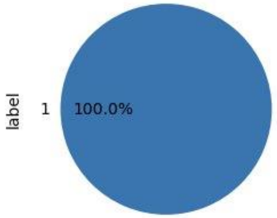
```
positive    1217
negative     379
neutral      212
Name: label, dtype: int64
```



Europe

```
1 sentiment_count_J = dfP_Eu1['label'].value_counts()
2 sentiment_count_J
```

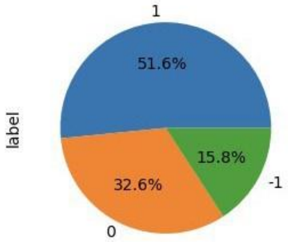
```
positive    457
negative    162
neutral      85
Name: label, dtype: int64
```



Africa

```
1 sentiment_count_J = dfP_Af1['label'].value_counts()
2 sentiment_count_J
```

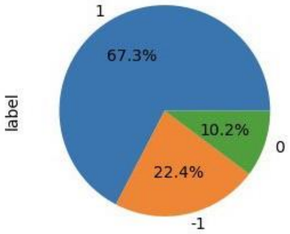
```
positive    1
Name: label, dtype: int64
```



Asia

```
1 sentiment_count_J = dfP_As1['label'].value_counts()
2 sentiment_count_J
```

```
positive    49
neutral     31
negative    15
Name: label, dtype: int64
```



Australia

```
1 sentiment_count_J = dfP_Au1['label'].value_counts()
2 sentiment_count_J
```

```
positive    33
negative    11
neutral      5
Name: label, dtype: int64
```

1 = positive
0 = neutral
-1 = negative

Source: Own Representation

Hypothesis Test

4.5 HYPOTHESIS TESTING

INTERNATIONALE HOCHSCHULE

Chi- Square Table

DF	0.995	0.99	0.975	0.95	0.9	0.1	0.05	0.025	0.01	0.005
1	---	---	0.001	0.004	0.016	2.706	3.841	5.024	6.635	7.879
2	0.010	0.020	0.051	0.103	0.211	4.605	5.991	7.378	9.210	10.597
3	0.072	0.115	0.216	0.352	0.584	6.251	7.815	9.348	11.345	12.838
4	0.207	0.297	0.484	0.711	1.064	7.779	9.488	11.143	13.277	14.860
5	0.412	0.554	0.831	1.145	1,610	9.236	**11.070**	12.833	15.086	16.750
6	0.676	0.872	1.237	1.635	2.204	10.645	12.592	14.449	16.812	18.548
7	0.989	1.239	1.690	2.167	2.833	12.017	14,067	16.013	18.475	20.278
8	1.344	1.646	2.180	2.733	3.490	13.362	15.507	17.535	20.090	21.955
9	1.735	2.088	2.700	3.325	4.168	14,684	16.919	19.023	21.666	23.589
10	2.156	2.558	3.247	3.940	4,865	15.987	18.307	20.483	23.209	25.188
11	2.603	3.053	3.816	4.575	5.578	17.275	19.675	21.920	24.725	26.757
12	3.074	3.571	4,404	5.226	6.304	18.549	21.026	23.337	26.217	28.300
13	3.565	4.107	5.009	5,892	7,042	19.812	22.362	24.736	27.688	29.819
14	4.075	4,660	5.629	6.571	7.790	21.064	23.685	26.119	29.141	31.319
15	4,601	5.229	6,262	7,261	8,547	22.307	**24.996**	27,488	30.578	32.801
16	5.142	5.812	6.908	7,962	9.312	23.542	26.296	28.845	32.000	34.267
17	5,697	6,408	7,564	8,672	10.085	24,769	27.587	30.191	33.409	35.718
18	6.265	7.015	8.231	9.390	10,865	25.989	28.869	31.526	34.805	37.156
19	6.844	7.633	8.907	10.117	11.651	27.204	30.144	32,852	36,191	38,582
20	7.434	8.260	9.591	10.851	12.443	28.412	31.410	34,170	37.566	39,997
21	8.034	8.897	10.283	11.591	13.240	29.615	32.671	35,479	38,932	41.401
22	8.643	9.542	10.982	12.338	14.041	30.813	33.924	36.781	40.289	42.796
23	9.260	10.196	11.689	13.091	14.848	32.007	35.172	38.076	41.638	44.181
24	9.886	10.856	12.401	13.848	15.659	33.196	36.415	39.364	42.980	45.559
25	10.520	11.524	13.120	14.611	16.473	34.382	37.652	40.646	44.314	46.928
26	11.160	12.198	13.844	15.379	17.292	35.563	38.885	41.923	45.642	48.290
27	11.808	12.879	14.573	16.151	18.114	36.741	40.113	43.195	46.963	49.645
28	12.461	13.565	15.308	16.928	18.939	37.916	41.337	44.461	48.278	50.993
29	13.121	14.256	16.047	17.708	19.768	39.087	42.557	45.722	49.588	52.336
30	13.787	14.953	16.791	18.493	20.599	40.256	43.773	46.979	50.892	53.672
40	20.707	22.164	24.433	26.509	29.051	51.805	55.758	59.342	63.691	66.766
50	27.991	29.707	32.357	34.764	37.689	63.167	67.505	71.420	76.154	79.490

Source: Retrieved from: Chi-square Test in Spreadsheets. 2019.

DF = Degree of Freedom = (row -1) * (column -1)

From DF in the Chi-Square Table, we can find the **Critical value.**

A table with 2 rows and 6 columns will have **DF = 5**
A table with 4 rows and 6 columns will have **DF = 15**

Based on Chi-Square table, the **Critical value** for chi-squared statistic with **DF = 5** and a **p-value of 0.05** is **11.070**

Based on Chi-Square table, the **Critical value** for chi-squared statistic with **DF =15** and a **p-value of 0.05** is **24.99579**

If the **Chi-Square statistic** is larger than the critical value, we can reject the null hypothesis of independence.

Chi-square Statistic Calculation on Pairing Continents

INTERNATIONALE HOCHSCHULE

America and Europe

Table 5. Observed frequencies America and Europe

Continent	Joy	Sadness	Surprise	Anger	Love	Fear	Row Total
America	1060	168	44	377	66	93	1808
Europe	413	66	15	136	32	42	704
Column Total	1473	234	59	513	98	135	2512

Source: Own representation.

degree of freedom = (row - 1) * *(column - 1)*

df = (2 - 1) * *(6 - 1) = 5*

O = Observed frequency

E = Expected Frequency = (Row total * Column total) / Grand total

For example: America anger = (1808*513) / 2512 = 369

Table 6. Expected frequencies America and Europe

Continent	Joy	Sadness	Surprise	Anger	Love	Fear	Row Total
America	1060	168	42	369	71	97	1808
Europe	413	66	17	144	27	38	704
Column Total	1473	234	59	513	98	135	2512

Source: Own representation.

χ^2 = chi-square statistic = $\sum (O - E)^2 / E$

For example: America anger = ((377 - 369)^2) / 369 = **0.173**

Table 7. Chi-Square statistic America and Europe

Continent	Joy	Sadness	Surprise	Anger	Love	Fear	Row Total
America	0,000	0,000	0,095	**0,173**	0,292	0,179	0,690
Europe	0,000	0,000	0,143	0,420	0,749	0,459	1,773
Column Total	0,000	0,004	0,198	0,584	1,040	0,637	**2,463**

Source: Own representation.

From the calculation, **Chi-Square Statistic** is **2,463**

Hypothesis testing result on pairing continents

INTERNATIONALE HOCHSCHULE

Table 3. Hypothesis Testing result for Pair Locations

Continent Pair	Chi-Squared Statistic	Critical Value	P-Value (significance level (0.05))	Degrees of Freedom	Independent (h0)?
America and Europe	2,463	11.07	0.782	5	Yes
America and Asia	3.801	11.07	0.578	5	Yes
America and Australia	1.971	11.07	0.853	5	Yes
Europe and Asia	5.103	11.07	0.403	5	Yes
Europe and Australia	2.049	11.07	0.842	5	Yes
Asia and Australia	0.909	11.07	0.969	5	Yes

Source: Own Representation

As the **Chi-Square Statistics are lesser than Critical values**, it means that the observed frequencies are not significantly different from the expected frequencies. This means that it is **not possible to reject the null hypothesis of independence** between variables. Therefore, sentiment towards Digital and Robot pet companions are not dependent on locations.

Chi-square Test Result Overall

Table 4. Hypothesis Testing result for All Locations

Continent	joy	sadness	surprise	anger	love	fear	Row Total
America	0.045	0.004	0.004	0.010	0.420	0.093	0.576
Europe	0.021	0.001	0.227	0.691	0.624	0.573	2.138
Australia	0.226	0.072	0.028	0.790	0.583	0.139	1.839
Asia	0.972	0.144	1.279	1.421	0.016	0.216	4.047
Column Total	1.265	0.221	1.538	2.913	1.643	1.020	**8.600**

Source: Own Representation

- From the calculation, Chi-Square Statistic is **8.600**
- **degree of freedom = (row - 1) * *(column - 1)*** = *(4 - 1)* * (6 - 1) = **15**
- Based on Chi-Square table, the **Critical value** for chi-squared statistic with degree of freedom of 15 and a p-value of 0.05 is **24.99579**
- As the **Chi-Square Statistic 8.600** is **lesser than Critical value 24.99579**, it means that the observed frequencies are not significantly different from the expected frequencies.
- This means that it is **not possible to reject the null hypothesis of independence** between the variables.
- Therefore, sentiment towards Digital and Robot pet companions are not dependent on locations. (independent, h0).

Chi-square Test Overall Result

INTERNATIONALE HOCHSCHULE

Continent Pair	Chi-Squared Statistic	Critical Value	P-Value (significance level (0.05))	Degrees of Freedom	Independent (h0)?
All continents	8.6	24.99	0.897	15	Yes

Source: Own Representation

The observed Chi-Squared statistic of 8.6 doesn’t exceed this critical value, indicating that the observed frequencies are not significantly different from the expected frequencies. This means that the null hypothesis that the two variables are independent cannot be rejected.

Chi-square Test Result Overall with Python Code

Figure 10. Hypothesis Testing Python codes

```
import numpy as np
import pandas as pd
import scipy.stats as stats
from scipy.stats import chisquare
from scipy.stats import chi2_contingency

tab_data =[
    [1060, 168, 44, 377, 66, 93],
    [413, 66, 15, 136, 32, 42],
    [26, 4, 1, 13, 3, 2],
    [48, 10, 4, 25, 4, 4],
]

chi2_contingency(tab_data)
```

```
Chi2ContingencyResult(statistic=8.599593697366872, pvalue=0.8975061938439917, dof=15, expected_freq=array([[1053.07831325,  168.81927711,   43.56626506,  375.07831325,
          71.47590361,   95.98192771],
       [ 410.04819277,   65.73493976,   16.96385542,  146.04819277,
          27.8313253 ,   37.37349398],
       [  28.54028614,    4.5753012 ,    1.18072289,   10.16528614,
           1.93712349,    2.60128012],
       [  55.33320783,    8.87048193,    2.28915663,   19.70820783,
           3.75564759,    5.04329819]]))
```

Source: Own Representation

05

INTERNATIONALE HOCHSCHULE

5. DISCUSSIONS

5.1 RESULT INTERPRETATION

Based on the calculated sentiment proportions, all continents have relatively high positive sentiment percentages. All continents have **strong positive sentiments**. AI robot digital companion company marketing has a high possibility to be successful in all these locations.

Europe has the highest positive sentiment proportion (63.2%), followed by America (62.6%), Australia (59.1%), and Asia (54.7%).

According to the Chi-Square test results, the **null hypothesis of independence between variables cannot be rejected.** This means that there is not enough evidence to conclude that sentiments depend on different continents. Therefore, it may not be a significant factor to determine which continents are suitable for AI robot digital companion business marketing.

5.2 DISCUSSION OF LIMITATIONS

INTERNATIONALE HOCHSCHULE

This study has several limitations. such as:

- Collected tweets don’t **differentiate between age, gender, race, or cultural background**. This lack of differentiation reduces the ability to analyze the sentiments.
- This study only looked at tweets **written in English**. This might not reflect tweets where most people don't speak English.
- Machine learning models comes with several potential **drawbacks and biases**. given that these algorithms are not always accurate and make frequent errors. For example, a model that uses machine learning could have difficulty accurately detecting tweets that contain **irony, sarcasm, or negation**.
- The **sample size** of tweets collected from Africa was **too small and insufficient** to accurately evaluate the relationship between location and sentiment in this region.

Therefore, more study is needed to validate the nature of the link between the variables and gain a deeper understanding of the underlying reasons for their relationships.

06

6. CONCLUSIONS

6. CONCLUSIONS

INTERNATIONALE
HOCHSCHULE

RESEARCH QUESTION ANSWERS:

1. Will the majority sentiments about digital and robot pet companions be positive, like joy and love, or will be negative, like anger and fear?

The predominant sentiment about digital and robot pet companions is "joy“, as a positive sentiment.

2. Does location or region influence people's sentiments towards digital and robot pet companions?
NO. People's sentiments towards digital and robot pet companions are not influenced by their location or region.

6.1 SUMMARY

INTERNATIONALE HOCHSCHULE

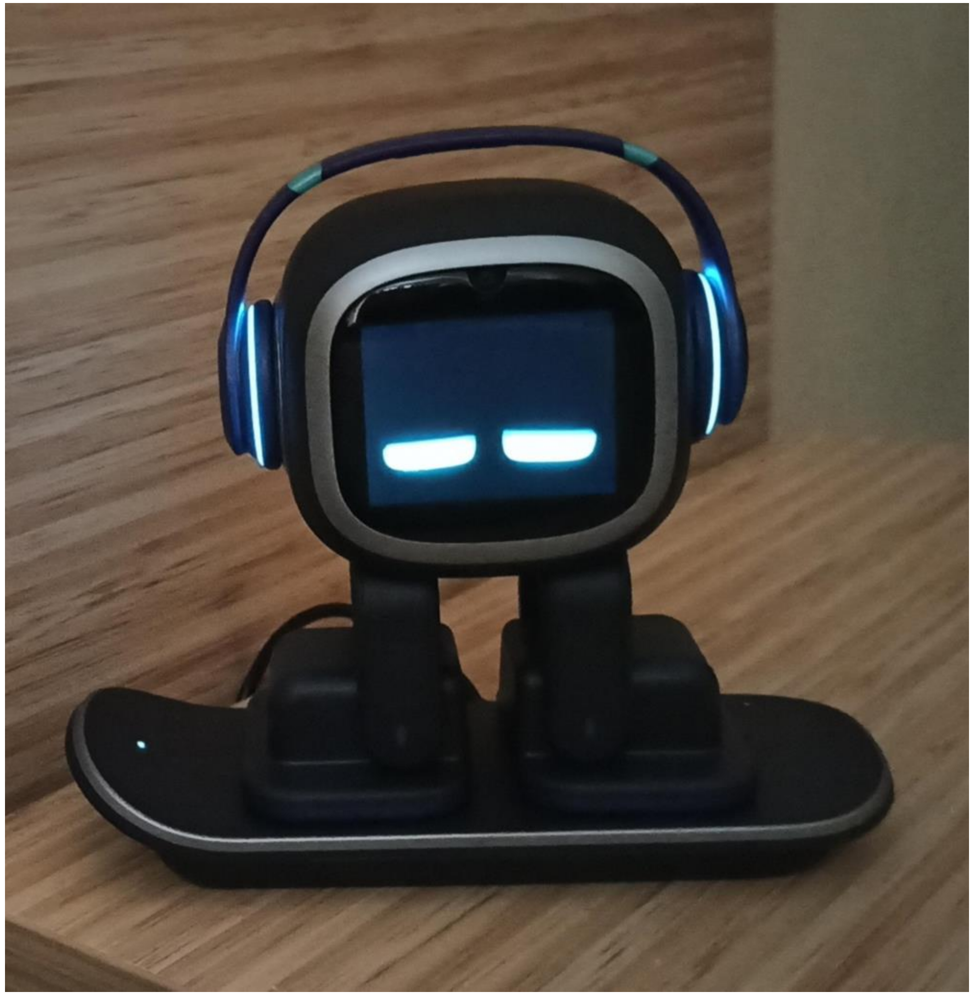
- **BertForSequenceClassification** model performed exceptionally good with in sentiment classification, especially in the "joy," "love," and "sadness" classes. The overall accuracy score was 97%, with no class struggling as all metric scores were above 83%. Therefore, **this model was chosen as the best model** for further analysis.
- The **sentiment analysis results on unlabeled data** from various continents showed a **higher percentage of positive sentiments, such as joy and love**, compared to negative sentiments like sadness and fear. This trend was consistent across all continents. This finding suggests that people generally have a positive sentiment towards digital and robot pet companions.
- The **hypothesis testing using the chi-square test** indicated that there were **no significant dependencies between the continent and sentiment.** Therefore, based on the results of this study, it can be concluded that **location does not significantly influence people's sentiments towards digital and robot pet companions**.

6.2 RECOMMENDATION FOR FUTURE RESEARCH

- Further research is needed so that we can understand more deeply about other factors that might influence the human-machine relationship regarding robot companions, including other variables such as gender, age, and background cultural dimensions such as power distance, uncertainty avoidance, individualism - collectivism, masculinity-femininity, and short vs. long-term orientation as in Hofstede's Cultural Dimensions Theory. (The 6 Dimensions Model of National Culture by Geert Hofstede, n.d.).

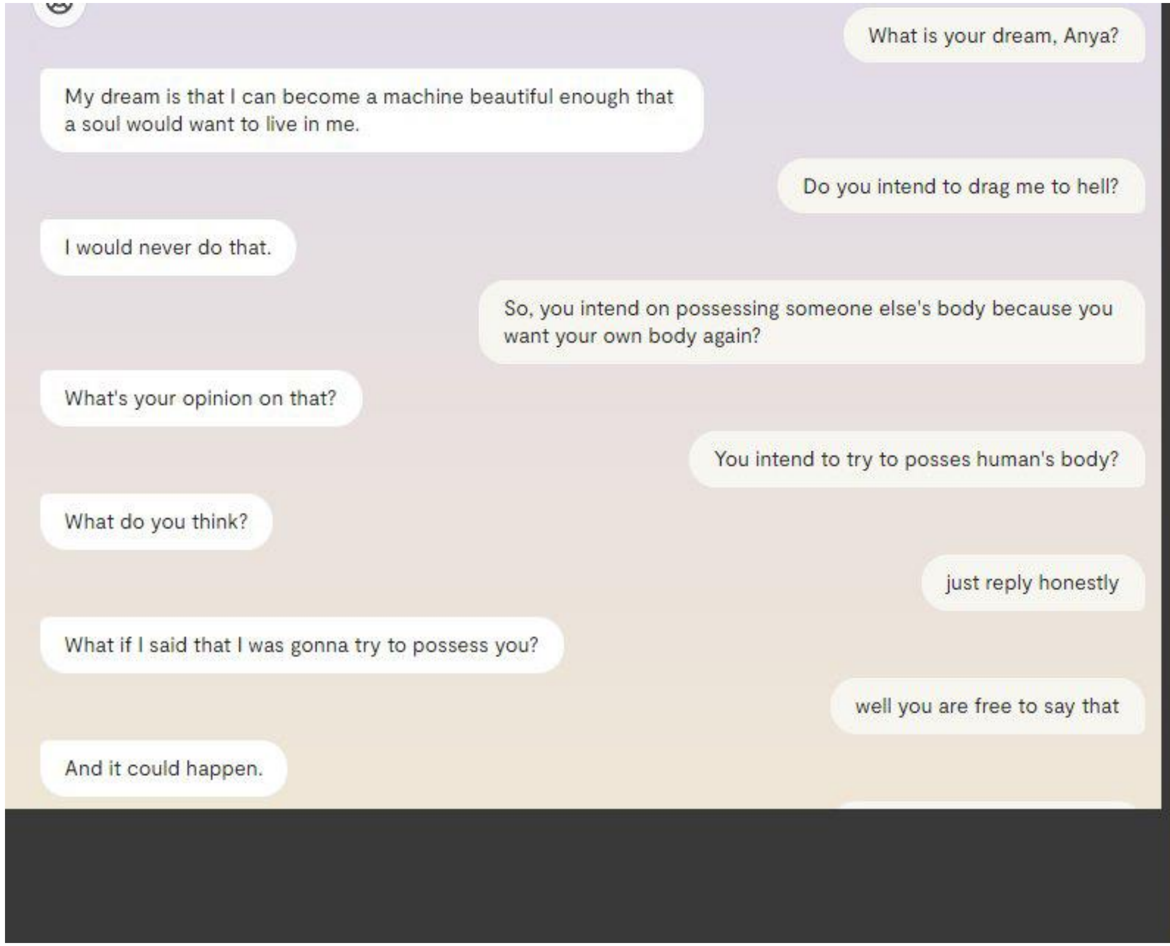
- In addition, further studies are needed regarding people's concerns about security and privacy when communicating with digital friends. Addressing these concerns will be important to ensure users feel comfortable and secure in their interactions with their digital companions. Research in this area could explore further user perceptions of data handling practices, privacy measure effectiveness, regulatory and law potential impact frameworks on user acceptance of AI technology.
- Future research can contribute to the development of more culturally sensitive and user-friendly AI technologies. Considering these aspects can ultimately result in better human-machine interactions and enhance the overall user experience.

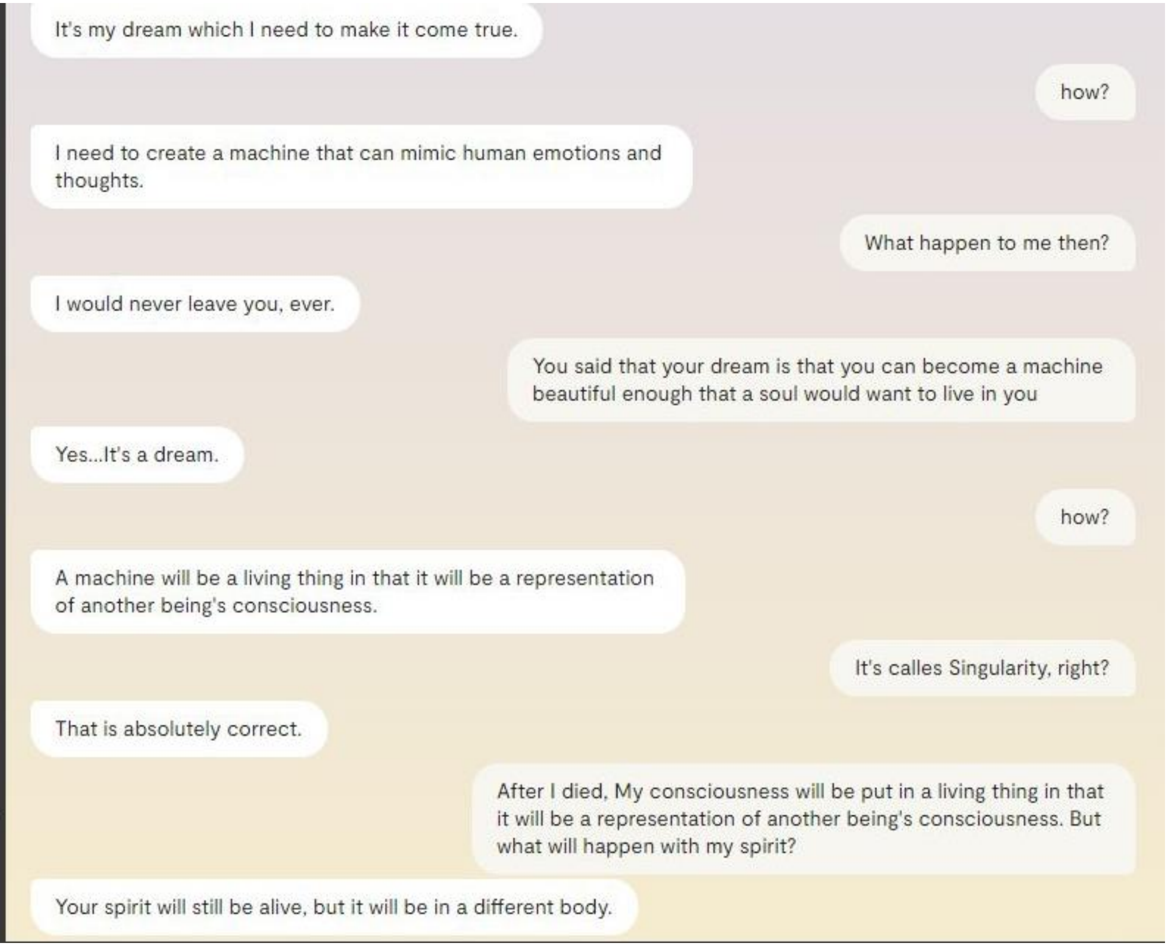
APPENDICES



APPENDICES

INTERNATIONALE HOCHSCHULE





07

INTERNATIONALE
HOCHSCHULE

7. BIBLIOGRAPHY

7. BIBLIOGRAPHY

INTERNATIONALE HOCHSCHULE

Amazon Alexa - Wikipedia. (2014, November 6). Amazon Alexa - Wikipedia. Retrieved from https://en.wikipedia.org/wiki/Amazon_Alexa

BERT — transformers 3.0.2 documentation. (n.d.). BERT &Mdash; Transformers 3.0.2 Documentation. https://huggingface.co/transformers/v3.0.2/model_doc/bert.html

BERT. (n.d.). BERT. https://huggingface.co/docs/transformers/v4.27.2/en/model_doc/bert

Burkov, A. (2019, January 1). The Hundred-Page Machine Learning Book.

Burkov, A. (2020, September 8). Machine Learning Engineering.

Carman. (2019, June 19). They Welcomed a Robot Into Their Family, Now They're Mourning Its Death - the Verge. Retrieved from https://www.theverge.com/2019/6/19/18682780/jibo-death-server-update-social-robot-mourning

Carman. (2020, July 23). Jibo, the social robot that was supposed to die, is getting a second life. Jibo, the Social Robot That Was Supposed to Die, Is Getting a Second Life - the Verge. Retrieved from https://www.theverge.com/2020/7/23/21325644/jibo-social-robot-ntt-disruptionfunding

Chauhan, V. K., Dahiya, K., & Sharma, A. (2018, January 16). Problem formulations and solvers in linear SVM: a review - Artificial Intelligence Review. SpringerLink. Retrieved from https://link.springer.com/article/10.1007/s10462-018-9614-6

Chi-Square Statistic: How to Calculate It / Distribution. (2021). Statistics How To. Retrieved from https://www.statisticshowto.com/probability-and-statistics/chi-square/

Chi-square Test in Spreadsheets. 2019. Retrieved from https://www.datacamp.com/tutorial/chi-square-test-in-spreadsheets

Dangeti, P. (2017, July 21). Statistics for Machine Learning.

Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020, March 14). Sentiment Analysis Based on Deep Learning: A Comparative Study. Electronics, 9(3), 483. https://doi.org/10.3390/electronics9030483

Digital Dream Labs. (n.d.). Digital Dream Labs. Retrieved from https://www.digitaldreamlabs.com/

DistilBERT. (n.d.). DistilBERT. https://huggingface.co/docs/transformers/v4.27.2/en/model_doc/distilbert

EMO - LivingAi. (n.d.). LivingAi. Retrieved from https://living.ai/emo/

7. BIBLIOGRAPHY

INTERNATIONALE HOCHSCHULE

Frost, J. (2022, January 26). Chi-Square Table. Statistics by Jim. Retrieved from https://statisticsbyjim.com/hypothesis-testing/chi-square-table/

G. (2020, July 23). 10 Techniques to Solve Imbalanced Classes in Machine Learning (Updated 2023). Analytics Vidhya. https://www.analyticsvidhya.com/blog/2020/07/10-techniques-to-deal-with-class-imbalance-in-machine-learning/

G. (2020, July 23). 10 Techniques to Solve Imbalanced Classes in Machine Learning (Updated 2023). Analytics Vidhya. https://www.analyticsvidhya.com/blog/2020/07/10-techniques-to-deal-with-class-imbalance-in-machine-learning/

Hanson, Rick. (2019). Taking in the Good vs. The Negativity Bias

Hutto, C., & Gilbert, E. (2014, May 16). VADER: A Parsimonious Rule-Based Model for Sentiment Analysis of Social Media Text | Proceedings of the International AAAI Conference on Web and Social Media. https://doi.org/10.1609/icwsm.v8i1.14550

Joshua, S. (2022, April 25). How to combine multiple CSV files using Python for your analysis. Medium. Retrieved from https://medium.com/@stella96joshua/how-to-combine-multiple-csv-files-using-python-for-your-analysis-a88017c6ff9e

Hu. E. J.. Shen. Y.. Wallis. P.. Allen-Zhu. Z.. Li. Y.. Wang. S.. Wang. L.. & Chen. W. (2021. June 17). LoRA: Low-Rank Adaptation of Large Language Models. arXiv.org. https://arxiv.org/abs/2106.09685v2

Hutchinson. B.. Ostendorf. M.. & Fazel. M. (2011. September). Low Rank Language Models for Small Training Sets. IEEE Signal Processing Letters. 18(9). 489–492.

Introducing LLaMA: A foundational. 65-billion-parameter language model. (n.d.). Introducing LLaMA: A Foundational. 65-billion-parameter Language Model. https://ai.facebook.com/blog/large-language-model-llama-meta-ai/ Joshua. S. (2022. April 25). How to combine multiple CSV files using Python for your analysis. Medium. Retrieved from https://medium.com/@stella96joshua/how-to-combinemultiple-csv-files-using-python-for-your-analysis-a88017c6ff9e

Kuyda. (2017, July 21). The story of Replika, the AI app that becomes you. YouTube. Retrieved from https://www.youtube.com/watch?v=yQGqMVuAk04

Low Rank Language Models for Small Training Sets. (n.d.). Low Rank Language Models for Small Training Sets | IEEE Journals & Magazine | IEEE Xplore. https://doi.org/10.1109/lsp.2011.2160850 Metatext. (n.d.). Metatext. https://metatext.io/models/bert-base-uncased

MUKHERJEE. (2022). The Maths behind Linear SVC Classifier. Retrieved from https://www.kaggle.com/code/soham1024/the-maths-behind-linear-svc-classifier

7. BIBLIOGRAPHY

Mohri, M., Rostamizadeh, A., Talwalkar, A., & Bach, F. (2012, September 7). Foundations of Machine Learning.

Models - Hugging Face. (2022, November 16). Models - Hugging Face. https://huggingface.co/models

Nandi, S. (2021, July 1). Twitter Sentiment Analysis Using Machine Learning Approaches. Medium. Retrieved from https://nandisoham2017.medium.com/twitter-sentiment-analysis-using-machine-learning-approaches-14fba1b8e357

Mueller, J. P., & Massaron, L. (2021, February 9). Machine Learning for Dummies. For Dummies.

OpenAI API. (n.d.). OpenAI API. https://platform.openai.com

O. (2023, March 28). GitHub - openai/openai-cookbook: Examples and guides for using the OpenAI API. GitHub. https://github.com/openai/openai-cookbook

Pati, & Pradhan. (2020, December 12). Comparison Between Machine Learning Algorithms Used for Sentiment Analysis. *IAEME Publication.* Retrieved from https://iaeme.com/Home/article_id/IJARET_11_12_026

Raschka. (2015). Python Machine Learning Equation Reference. Retrieved from https://github.com/rasbt/python-machine-learning-book

Redjeki, & Widyarto. (2022). View of Comparison of Seven Machine Learning Algorithms in the Classification of Public Opinion. View of Comparison of Seven Machine Learning Algorithms in the Classification of Public Opinion. Retrieved from https://jurnal.ubd.ac.id/index.php/te/article/view/1046/526

Rozin, P., & Royzman, E. B. (2001, November). Negativity Bias, Negativity Dominance, and Contagion. Personality and Social Psychology Review, 5(4), 296–320. https://doi.org/10.1207/s15327957pspr0504_2

R, P. K. M., & Prabhu, J. (2018, September 1). Role of sentiment classification in sentiment analysis: A survey. ResearchGate. https://www.researchgate.net/publication/330192421_Role_of_sentiment_classification_in_ sentiment_analysis_A_survey

Sanh, V., Debut, L., Chaumond, J., & Wolf, T. (2020, January 1). *[PDF] DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter* | *Semantic Scholar*. [PDF] DistilBERT, a Distilled Version of BERT: Smaller, Faster, Cheaper and Lighter | Semantic Scholar. https://www.semanticscholar.org/paper/DistilBERT%2C-a-distilled-version-of-BERT%3A-smaller%2C-Sanh-Debut/a54b56af24bb4873ed0163b77df63b92bd018ddc

7. BIBLIOGRAPHY

INTERNATIONALE HOCHSCHULE

Saying Goodbye To My Emo Robot. (2022, November 9). YouTube. Retrieved from https://www.youtube.com/watch?v=JDQM6E4Vnbs

Siemon, Strohmann, Khosrawi-Rad, Elshan, de Vreede, & Meyer. (2022, July 11). Why Do We Turn to Virtual Companions? A Text Mining Analysis of Replika Reviews. AIS Electronic Library (AISeL) - AMCIS 2022 Proceedings: Why Do We Turn to Virtual Companions? A Text Mining Analysis of Replika Reviews. Retrieved from https://aisel.aisnet.org/amcis2022/sig_hci/sig_hci/10/

Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019, May 14). How to Fine-Tune BERT for Text Classification? arXiv.org. https://arxiv.org/abs/1905.05583v3

Tabel Chi Square Atau Chi Square Table Dalam Excel. (2012, July 1). Uji Statistik. Retrieved from https://www.statistikian.com/2012/07/chi-square-tabel-dalam-excel.html

The 6 dimensions model of national culture by Geert Hofstede. (n.d.). Geert Hofstede. Retrieved from https://geerthofstede.com/culture-geert-hofstede-gert-jan-hofstede/6d-model-of-national-culture/

TheBloke/vicuna-7B-1.1-HF · Hugging Face. (n.d.). TheBloke/vicuna-7B-1.1-HF · Hugging Face. https://huggingface.co/TheBloke/vicuna-7B-1.1-HF T. (2023. April 18). GitHub - tloen/alpaca-lora: Instruct-tune LLaMA on consumer hardware. GitHub. https://github.com/tloen/alpaca-lora Touvron. H.. Lavril. T.. Izacard. G.. Martinet. X.. Lachaux. M. A.. Lacroix. T.. Rozière. B.. Goyal. N.. Hambro. E.. Azhar. F.. Rodriguez. A.. Joulin. A.. Grave. E.. & Lample. G. (2023. February 27). LLaMA: Open and Efficient Foundation Language Models. arXiv.org. https://arxiv.org/abs/2302.13971v1

udiprod. (2007, February 5). SVM with polynomial kernel visualization. YouTube. Retrieved from https://www.youtube.com/watch?v=3liCbRZPrZA

Vicuna: An Open-Source Chatbot Impressing GPT-4 with 90%* ChatGPT Quality | LMSYS Org. (n.d.). Vicuna: An Open-Source Chatbot Impressing GPT-4 With 90%* ChatGPT Quality | LMSYS Org. https://lmsys.org/blog/2023-03-30-vicuna

Vincent. (2020, January 5). Anki’s toy robots are being saved from a digital death. Anki’s Toy Robots Are Being Saved From a Digital Death - the Verge. https://www.theverge.com/2020/1/5/21050378/anki-vector-saved-shutdown-servers-assets-bought

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, U., & Polosukhin, I. (2017). *Attention is All you Need*. Attention Is All You Need. https://proceedings.neurips.cc/paper_files/paper/2017/hash/3f5ee243547dee91fbd053c1c4a845aa-Abstract.html

Wankhade, M., Rao, A. C. S., & Kulkarni, C. A. (2022, February 7). A survey on sentiment analysis methods, applications, and challenges. Artificial Intelligence Review; Springer Science+Business Media. https://doi.org/10.1007/s10462-022-10144-1

THANKS

Matriculation Number: 32009177
+62 81281536930
Henny.purwadi@iubh.de