# BAB 4

**HASIL PENELITIAN**



## *Testing Environment*

### Spesifikasi Sistem

#### Device

Penelitian yang diimplementasikan ke dalam sebuah web aplikasi dibangun dan dijalankan pada perangkat pc dengan spesifikasi sebagai berikut.

|  |  |  |
| --- | --- | --- |
| Operating System | : | Windows 10 Pro Edition |
| CPU | : | Intel Core i7-4710HQ CPU @2.50 GHz |
| RAM | : | 8.00 GB |
| Internal Memory | : | 1 TB |



#### 

#### Server

Sistem web aplikasi ini dijalankan dengan Apache HTTP Server 2.4.

### Perangkat Lunak

Berikut adalah perangkat lunak yang digunakan dalam penelitian ini

1. Eclipse Neon 4.6 IDE
2. Notepad++ 7.3.3
3. Microsoft Excel 2010
4. XAMPP 3.2.2
5. Apache HTTP Server 2.4

### *Library*

Berikut adalah *library* yang digunakan dalam pembangunan sistem prediksi.

1. Scikit-learn 0.18.1
2. LIWC2015
3. SPLICE 0.9.0
4. Numpy 1.13.0
5. Pandas 0.20.0
6. Keras 2.0.4
7. Theano 0.9.0
8. Imbalanced-learn 0.2.1
9. NLTK 3.2.4

## Hasil

### Distribusi Data

Distribusi data dari *dataset* my*Personality* dapat dilihat pada Tabel 4.1 dengan total sebanyak 250 data *user* Facebook.

**Tabel 4.1** Distribusi data my*Personality*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Value** | ***Openness*** | ***Conscientiousness*** | ***Extraversion*** | ***Agreeableness*** | ***Neuroticism*** |
| Yes | 176 | 130 | 96 | 134 | 99 |
| No | 74 | 120 | 154 | 116 | 151 |

Distribusi data dari *dataset* *Manual gathering* dapat dilihat pada Tabel 4.2 dengan total sebanyak 150 data *user* Facebook yang didapatkan secara manual.

**Tabel 4.2** Distribusi data *Manual gathering*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Value** | ***Openness*** | ***Conscientiousness*** | ***Extraversion*** | ***Agreeableness*** | ***Neuroticism*** |
| Yes | 97 | 63 | 38 | 81 | 50 |
| No | 53 | 87 | 112 | 69 | 100 |

Distribusi data dari gabungan kedua *dataset* di atas dapat dilihat pada Tabel 4.3 dengan total sebanyak 400 data *user*.

**Tabel 4.3** Distribusi data gabungan my*Personality* dan Manual Data Gathering

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Value** | ***Openness*** | ***Conscientiousness*** | ***Extraversion*** | ***Agreeableness*** | ***Neuroticism*** |
| Yes | 273 | 193 | 134 | 215 | 149 |
| No | 127 | 207 | 266 | 185 | 251 |

### Faktor Skenario Percobaan

Beberapa faktor skenario percobaan yang akan digunakan selama proses *testing* untuk mendapatkan hasil terbaik dari implementasi *machine learning* ini terbagi dua yaitu:

* + - 1. ***Features* *Selection***

Feature utama yang digunakan dari penelitian ini yaitu LIWC sebanyak 85 *features*, SPLICE sebanyak 94 *features*, dan SNA *features* sebanyak 7 *features* akan melewati proses *feature selection*s atau pemilihan fitur selama proses *testing* berjalan untuk mendapatkan feature yang paling optimal dalam menghasilkan tingkat akurasi paling tinggi.

Modul *Features selection* ini berjalan secara otomatis dan dinamis selama proses berjalan dan feature paling optimal akan muncul di akhir bersamaan dengan tingkat akurasinya. *Features selection* hanya akan digunakan untuk skenario implementasi *machine learning*.

* + - 1. ***Resampling***

*Resampling* adalah proses di mana *dataset* yang ada dilakukan manipulasi data dan duplikasi data untuk membuat sebuah data memiliki elemen yang seimbang. Proses *resampling* terbagi menjadi 2 yaitu:

1. *Oversampling*

Menambahkan data ke data dengan persentase minoritas. Tujuan *Oversampling* pada penelitian ini karena ada *dataset* di mana elemennya tidak seimbang. Sebagai contoh terdapat 70% pria dan 30% wanita. Hasil dari *testing* biasanya akan buruk dikarenakan faktor tersebut. Untuk itu, *oversampling* akan menambahkan elemen yang menjadi minoritas agar data menjadi seimbang.

1. *Undersampling*

Proses *undersampling* hanya berbanding terbalik dengan *Oversampling*. Jika *Oversampling* menambahkan data, *Undersampling* mengurangi data yang memiliki persentase mayoritas untuk menyeimbangkan data.

Faktor skenario *resampling* ini akan digunakan pada skenario implementsi *machine learning* dan juga skenario implementasi *deep learning*.

### Implementasi *Machine learning*

* + - 1. **Skenario**

Implementasi *machine learning* dalam *testing* ini akan menggunakan 5 jenis algoritma (Support Vector Machine, Naïve Bayes, Logistic Regression, Gradient Boosting, dan LDA) dan 3 *features* utama (LIWC, SPLICE, dan SNA). Kemudian, *dataset* yang digunakan terbagi menjadi 3 seperti dijelaskan pada Bagian 4.2.1 mengenai distribusi data.

Dari faktor skenario diatas pada Bagian 4.2.2, *machine learning* akan menggunakan kedua faktor tersebut yaitu *Features selection* (Bagian 4.2.2.1) dan *Resampling* (Bagian 4.2.2.2) sehingga dapat dibentuk skenario percobaan untuk *Machine learning* sebagai berikut:

* + 1. Penggunaan *dataset* my*Personality*, percobaan tanpa menggunakan *Features selection* dan tanpa menggunakan *Resampling*.
    2. Penggunaan *dataset* my*Personality*, percobaan dengan menggunakan *Features selection*.
    3. Penggunaan *dataset* my*Personality*, percobaan dengan menggunakan *Undersampling*.
    4. Penggunaan *dataset* my*Personality*, percobaan dengan menggunakan *Oversampling*.
    5. Penggunaan *dataset* my*Personality*, percobaan dengan menggunakan *Features selection* dan dengan menggunakan *Undersampling*.
    6. Penggunaan *dataset* my*Personality*, percobaan dengan menggunakan *Features selection* dan dengan menggunakan *Oversampling*.
    7. Penggunaan *dataset* *Manual gathering*, percobaan tanpa menggunakan *Features selection* dan tanpa menggunakan *Resampling*.
    8. Penggunaan *dataset* *Manual gathering*, percobaan dengan menggunakan *Features selection*.
    9. Penggunaan *dataset* *Manual gathering*, percobaan dengan menggunakan *Undersampling*.
    10. Penggunaan *dataset* *Manual gathering*, percobaan dengan menggunakan *Oversampling*.
    11. Penggunaan *dataset* *Manual gathering*, percobaan dengan menggunakan *Features selection* dan dengan menggunakan *Undersampling*.
    12. Penggunaan *dataset* *Manual gathering*, percobaan dengan menggunakan *Features selection* dan dengan menggunakan *Oversampling*.
    13. Penggunaan *dataset* gabungan, percobaan tanpa menggunakan *Features selection* dan tanpa menggunakan *Resampling*.
    14. Penggunaan *dataset* gabungan, percobaan dengan menggunakan *Features selection*.
    15. Penggunaan *dataset* gabungan, percobaan dengan menggunakan *Undersampling*.
    16. Penggunaan *dataset* gabungan, percobaan dengan menggunakan *Oversampling*.
    17. Penggunaan *dataset* gabungan, percobaan dengan menggunakan *Features selection* dan dengan menggunakan *Undersampling*.
    18. Penggunaan *dataset* gabungan, percobaan dengan menggunakan *Features selection* dan dengan menggunakan *Oversampling*.

Keterangan pada tabel percobaan:

|  |  |
| --- | --- |
|  | = *Header* Tabel percobaan yang menggunakan *dataset* my*Personality* |
|  | = *Header* Tabel percobaan yang menggunakan *dataset* *Manual gathering* |
|  | = *Header* Tabel percobaan yang menggunakan *dataset* gabungan |
|  |  |
| O | = *Traits* *Openness* |
| C | = *Traits* *Conscientiousness* |
| E | = *Traits* *Extraversion* |
| A | = *Traits* *Agreeableness* |
| N | = *Traits* *Neuroticism* |
| **Bold** | = Akurasi tertinggi pada masing-masing *Features* |
| **Highlight** | = Akurasi tertinggi pada sebuah *Traits* |

* + - 1. ***Dataset* my*Personality***

Percobaan dengan skenario 1-6 dilakukan dengan menggunakan *dataset* my*Personality* (Kosinski, 2015). Untuk kemudian dibandingkan dengan penggunaan *dataset* yang didapatkan secara manual dan gabungan keduanya.

* + - * 1. **Skenario 1 (Percobaan tanpa menggunakan *Features selection* dan tanpa *Resampling*)**

**Tabel 4.4** Hasil Percobaan Skenario 1 tanpa *Features selection* dan tanpa *Resampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 64.80% | 50.00% | 51.60% | **54.80%** | 54.40% |
| SVM | 70.00% | 50.00% | 60.00% | 46.80% | 58.00% |
| Logistic Regression | **70.40%** | 50.00% | 61.60% | 51.20% | 60.40% |
| GRADIENT BOOSTING | 63.20% | 53.20% | 57.60% | 53.60% | 56.80% |
| LDA | 62.00% | **54.00%** | **64.40%** | 53.60% | **60.80%** |
| SPLICE | NAÏVE BAYES | 59.20% | **52.00%** | 53.60% | 53.60% | 47.20% |
| SVM | 62.00% | 51.60% | 53.20% | 54.00% | 50.40% |
| Logistic Regression | 63.60% | 51.60% | **58.00%** | 51.60% | 54.40% |
| GRADIENT BOOSTING | 62.00% | 50.00% | 54.40% | 50.80% | 54.00% |
| LDA | **64.00%** | 51.60% | 54.00% | **58.00%** | **55.20%** |
| SNA | NAÏVE BAYES | 58.00% | 54.80% | **68.80%** | 45.60% | 52.40% |
| SVM | 58.80% | 53.60% | 51.60% | 42.00% | 38.80% |
| Logistic Regression | **70.00%** | 54.00% | 68.40% | 52.00% | **58.40%** |
| GRADIENT BOOSTING | 63.20% | 48.80% | 68.00% | **55.20%** | 54.00% |
| LDA | 69.20% | **56.40%** | 66.00% | 54.80% | 57.60% |
| LIWC + SPLICE | NAÏVE BAYES | 64.80% | 50.80% | 53.60% | **54.00%** | 46.80% |
| SVM | 59.60% | 50.80% | 54.00% | 51.20% | 50.80% |
| Logistic Regression | 63.60% | 52.40% | 56.80% | 51.20% | 54.00% |
| GRADIENT BOOSTING | **65.60%** | **54.40%** | 54.80% | 47.60% | **60.00%** |
| LDA | 58.00% | 54.00% | **63.20%** | 49.60% | 50.00% |
| LIWC + SNA | NAÏVE BAYES | 58.00% | **55.20%** | **68.80%** | 45.60% | 52.40% |
| SVM | 58.80% | 53.60% | 51.60% | 41.20% | 38.80% |
| Logistic Regression | **70.00%** | 54.00% | 68.40% | 51.20% | 59.20% |
| GRADIENT BOOSTING | 64.80% | 50.40% | 60.00% | **55.20%** | **60.40%** |
| LDA | 62.40% | 53.60% | 66.40% | 52.80% | 59.60% |
| SPLICE + SNA | NAÏVE BAYES | 40.00% | 49.60% | **64.00%** | 51.20% | 44.00% |
| SVM | 62.80% | 50.80% | 63.20% | **54.00%** | 47.20% |
| Logistic Regression | **68.40%** | **54.80%** | 63.60% | 45.20% | **59.20%** |
| GRADIENT BOOSTING | 64.80% | 50.00% | 62.40% | 48.40% | 55.60% |
| LDA | 60.80% | 53.20% | 56.40% | 53.20% | 56.40% |
| LIWC + SPLICE + SNA | NAÏVE BAYES | 40.00% | 49.60% | 64.00% | 51.20% | 44.00% |
| SVM | 62.80% | 50.80% | 63.20% | 54.00% | 47.20% |
| Logistic Regression | **68.40%** | **56.40%** | **65.60%** | 46.80% | 57.60% |
| GRADIENT BOOSTING | 63.60% | 53.60% | 60.80% | **52.80%** | **60.00%** |
| LDA | 60.40% | 54.40% | 63.20% | 47.60% | 52.40% |

Tabel 4.4 diatas merupakan percobaan yang menggunakan *dataset* myPersonality dan 7 fitur berbeda. Evaluasi testing dilakukan dengan 10-fold cross validation. Percobaan skenario 1 ini tidak menggunakan proses *Features selection* dan juga proses Resampling

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 70.40% dengan menggunakan algoritma Logistic Regression dan fitur LIWC. Hasil akurasi tertinggi kedua adalah 68.80% dengan menggunakan algoritma Naïve Bayes, fitur LIWC+SNA dan algoritma Naïve Bayes, fitur SNA. Hasil tertinggi yang didapatkan menunjukkan dominasi fitur LIWC yang kemungkinan besar memang tepat untuk diterapkan pada sistem prediksi ini.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Openness* yang menggunakan algoritma Logistic Regression dan fitur LIWC.

* + - * 1. **Skenario 2 (Percobaan dengan menggunakan *Features selection*)**

**Tabel 4.5** Hasil Percobaan Skenario 2 dengan *Features selection*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 70.00% | 48.80% | 59.60% | 47.60% | 51.20% |
| SVM | **70.40%** | **56.00%** | 61.60% | 52.40% | **60.40%** |
| Logistic Regression | **70.40%** | 53.60% | 61.60% | 53.60% | **60.40%** |
| GRADIENT BOOSTING | 62.00% | 54.40% | 60.80% | **56.40%** | 58.80% |
| LDA | 69.60% | 52.80% | **62.80%** | 53.60% | **60.40%** |
| SPLICE | NAÏVE BAYES | 59.20% | 51.60% | 53.60% | **54.00%** | 49.60% |
| SVM | 44.00% | 48.40% | 51.60% | 51.20% | 54.80% |
| Logistic Regression | **65.60%** | 52.00% | **56.40%** | 52.00% | **57.60%** |
| GRADIENT BOOSTING | 63.60% | 50.80% | 54.40% | 51.20% | 53.60% |
| LDA | 62.40% | **54.00%** | 53.60% | **54.00%** | **57.60%** |
| SNA | NAÏVE BAYES | 58.00% | **54.80%** | **68.80%** | 45.60% | 52.40% |
| SVM | 58.80% | 53.60% | 51.60% | 41.60% | 38.80% |
| Logistic Regression | **70.00%** | 53.60% | 68.40% | 50.80% | 57.60% |
| GRADIENT BOOSTING | 61.20% | 48.40% | 68.00% | 50.40% | **59.20%** |
| LDA | **70.00%** | 52.40% | 68.00% | **56.80%** | 58.40% |
| LIWC + SPLICE | NAÏVE BAYES | 60.80% | 52.00% | 52.80% | **53.60%** | 48.80% |
| SVM | 45.20% | 51.60% | 54.00% | 50.40% | 54.00% |
| Logistic Regression | **65.20%** | 52.00% | 56.40% | 51.60% | **56.80%** |
| GRADIENT BOOSTING | 63.20% | 52.00% | **58.00%** | 50.80% | 53.60% |
| LDA | 62.00% | **55.20%** | 55.60% | 52.80% | 56.40% |
| LIWC + SNA | NAÏVE BAYES | 58.00% | **54.80%** | **68.80%** | 45.60% | 52.40% |
| SVM | 58.80% | 53.60% | 51.60% | 41.60% | 38.80% |
| Logistic Regression | **70.00%** | 54.00% | 68.40% | 52.00% | 58.00% |
| GRADIENT BOOSTING | 62.40% | 56.00% | 66.80% | **60.40%** | **57.20%** |
| LDA | 68.80% | 51.60% | 68.40% | 54.80% | 56.40% |
| SPLICE + SNA | NAÏVE BAYES | 40.40% | 50.00% | 64.80% | 52.00% | 44.40% |
| SVM | 62.80% | 50.80% | 63.20% | **56.80%** | 46.40% |
| Logistic Regression | **66.40%** | **54.80%** | **65.20%** | 44.80% | 58.00% |
| GRADIENT BOOSTING | 65.20% | 50.80% | 61.60% | 52.80% | 54.00% |
| LDA | 64.00% | 52.80% | 60.40% | 54.00% | **59.20%** |
| LIWC + SPLICE + SNA | NAÏVE BAYES | 40.40% | 50.00% | **64.80%** | 52.00% | 44.00% |
| SVM | 62.80% | 50.80% | 63.20% | **56.80%** | 47.20% |
| Logistic Regression | **68.80%** | 54.80% | 63.60% | 46.00% | 57.60% |
| GRADIENT BOOSTING | 61.60% | 50.00% | 60.40% | 50.00% | **60.00%** |
| LDA | 61.60% | **56.40%** | 60.00% | 53.60% | 52.40% |

Tabel 4.5 diatas merupakan percobaan yang menggunakan *dataset* myPersonality dan 7 fitur berbeda. Percobaan skenario 2 ini hanya menggunakan proses *Features selection* yaitu melakukan pemilihan terhadap fitur-fitur yang memiliki korelasi tinggi dan dianggap optimal terhadap peningkatan hasil akurasi*.* Evaluasi testing dilakukan dengan 10-fold cross validation.

Dari Tabel 4.5 diatas, dapat dilihat bahwa di skenario 2 ini nilai akurasi tertinggi yaitu 70.40% dengan algoritma Logistic Regression dan fitur LIWC. Hasil ini persis sama dengan hasil yang ada pada skenario 1. Hasil tertinggi kedua dengan nilai 68.80 dan menggunakan algoritma Naïve Bayes. Hasil akurasi tertinggi kedua ini juga persis sama dengan hasil yang didapatkan pada skenario 1, Tabel 4.4. *Traits* yang memiliki akurasi tertinggi pada skenario ini adalah *traits* *Openness*.

Kesimpulan skenario ini bahwa *Features selection* tidak terlalu berpengaruh terhadap hasil akurasi karena dapat dilihat beberapa akurasi masih mempunyai nilai yang sama dengan skenario 1. Hanya pada *traits* *Agreeableness*, akurasi mengalami peningkatan dari percobaan di skenario sebelumnya. Akurasi tertinggi *Agreeableness* pada skenario ini naik 2.40% dari hasil pada Tabel 4.4 menjadi 60.40%.

* + - * 1. **Skenario 3 (Percobaan dengan menggunakan *Undersampling*)**

**Tabel 4.6** Hasil Percobaan Skenario 3 dengan *Undersampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 35.60% | 50.00% | 46.80% | **55.60%** | 49.20% |
| SVM | 44.00% | 51.60% | 53.60% | 46.80% | 48.40% |
| Logistic Regression | 43.20% | **54.40%** | 49.60% | 48.40% | 52.40% |
| GRADIENT BOOSTING | 44.00% | 51.60% | 55.60% | 52.80% | 53.60% |
| LDA | **49.60%** | 54.00% | **56.00%** | 52.40% | **58.00%** |
| SPLICE | NAÏVE BAYES | 37.60% | 50.00% | 42.40% | **54.00%** | 43.60% |
| SVM | 39.60% | 51.60% | 48.00% | 50.40% | 50.80% |
| Logistic Regression | 45.20% | 50.80% | **51.60%** | 50.80% | 50.40% |
| GRADIENT BOOSTING | 43.20% | 49.20% | 44.80% | 49.60% | **54.00%** |
| LDA | **53.60%** | **52.40%** | **51.60%** | 52.80% | 52.80% |
| SNA | NAÏVE BAYES | 38.00% | 57.60% | **68.00%** | 48.00% | 50.80% |
| SVM | **46.40%** | 51.20% | 54.80% | 40.00% | 42.00% |
| Logistic Regression | 41.20% | 52.00% | 66.40% | 52.80% | 58.00% |
| GRADIENT BOOSTING | 41.60% | 53.20% | 55.20% | **56.40%** | 53.60% |
| LDA | **46.40%** | **58.40%** | 56.40% | 53.60% | **60.00%** |
| LIWC + SPLICE | NAÏVE BAYES | 41.20% | 50.40% | 45.60% | **55.60%** | 45.60% |
| SVM | 44.00% | 53.60% | 46.00% | 41.60% | **54.40%** |
| Logistic Regression | 45.20% | 51.60% | 49.60% | 49.60% | 50.00% |
| GRADIENT BOOSTING | 46.80% | **55.20%** | 48.80% | 50.00% | 51.60% |
| LDA | **52.00%** | 52.40% | **57.60%** | 50.40% | 54.00% |
| LIWC + SNA | NAÏVE BAYES | 38.00% | **57.60%** | **68.00%** | 48.00% | 50.80% |
| SVM | 47.20% | 57.20% | 61.20% | 42.40% | 45.60% |
| Logistic Regression | 42.00% | 52.00% | 66.40% | 52.80% | **58.40%** |
| GRADIENT BOOSTING | 48.40% | 50.80% | 51.20% | 50.40% | 57.60% |
| LDA | **48.80%** | 54.80% | 56.80% | **54.00%** | 50.80% |
| SPLICE + SNA | NAÏVE BAYES | 40.80% | 51.60% | 52.80% | 52.40% | 42.00% |
| SVM | 44.80% | 54.00% | 52.00% | 50.40% | 52.00% |
| Logistic Regression | 48.40% | **55.20%** | **60.00%** | 50.80% | 56.40% |
| GRADIENT BOOSTING | 47.60% | 51.20% | 55.60% | 53.20% | **56.80%** |
| LDA | **53.20%** | 54.80% | 50.80% | **54.80%** | 53.60% |
| LIWC + SPLICE + SNA | NAÏVE BAYES | 42.40% | 51.60% | 52.80% | 52.40% | 42.00% |
| SVM | 44.40% | 49.60% | 54.80% | 51.60% | 50.00% |
| Logistic Regression | 45.20% | 51.60% | **58.80%** | **55.60%** | 54.00% |
| GRADIENT BOOSTING | 50.00% | 49.60% | 55.60% | 49.60% | **56.80%** |
| LDA | **54.40%** | **54.80%** | 54.00% | 48.80% | 46.80% |

Tabel 4.6 diatas merupakan percobaan yang menggunakan *dataset* myPersonality dan 7 fitur berbeda. Percobaan skenario 3 ini menggunakan proses *Undersampling* yaitu menghilangkan data mayoritas dalam distribusi *dataset*. Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 68.00% dengan menggunakan algoritma Naïve Bayes, fitur LIWC dan fitur LIWC+SNA. Hasil akurasi tertinggi kedua adalah 60.00% dengan menggunakan algoritma Linear Discriminant Analysis dan fitur SNA. Hasil skenario ini menunjukkan penggunaan *Undersampling* pada *dataset* myPersonality ini tidak berhasil meningkatkan hasil akurasi secara signifikan.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion* dengan menggunakan algoritma Naïve Bayes. Hasil yang didapatkan pada skenario ini menunjukkan hanya satu *traits* yang mengalami peningkatan akurasi yaitu *traits* *Conscientiousness* dengan 58.40, meningkat 2% dari hasil akurasi tertinggi untuk *traits* ini pada skenario sebelumnya.

* + - * 1. **Skenario 4 (Percobaan dengan menggunakan *Oversampling*)**

**Tabel 4.7** Hasil Percobaan Skenario 4 dengan *Oversampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 48.80% | 49.20% | 50.00% | 52.00% | 50.80% |
| SVM | 54.80% | 50.40% | 52.80% | 49.20% | 53.20% |
| Logistic Regression | 56.00% | 53.20% | 50.80% | 48.40% | 53.60% |
| GRADIENT BOOSTING | **57.20%** | **56.00%** | 57.20% | 51.60% | 57.60% |
| LDA | 55.20% | 54.80% | **60.00%** | **53.20%** | **58.80%** |
| SPLICE | NAÏVE BAYES | 41.20% | **52.00%** | 52.00% | **56.40%** | 43.20% |
| SVM | 49.60% | 49.60% | 49.60% | 48.40% | 54.00% |
| Logistic Regression | 54.00% | 51.20% | **54.80%** | 51.60% | 48.80% |
| GRADIENT BOOSTING | **60.80%** | 50.40% | 51.60% | 54.40% | **56.80%** |
| LDA | 57.60% | **52.00%** | 52.00% | **56.40%** | 55.60% |
| SNA | NAÏVE BAYES | 37.20% | **59.20%** | **65.60%** | 44.00% | 49.60% |
| SVM | 50.40% | 55.20% | 47.20% | 45.20% | 38.40% |
| Logistic Regression | 45.20% | 54.40% | 61.20% | 49.20% | **58.00%** |
| GRADIENT BOOSTING | **57.20%** | 48.80% | 64.80% | **52.80%** | 55.20% |
| LDA | 52.40% | 58.40% | 59.60% | 52.00% | 57.20% |
| LIWC + SPLICE | NAÏVE BAYES | 46.40% | 50.80% | 51.20% | **56.40%** | 41.60% |
| SVM | 50.00% | 52.00% | 52.40% | 46.00% | 51.60% |
| Logistic Regression | 54.80% | 50.00% | 57.20% | 51.20% | 49.60% |
| GRADIENT BOOSTING | **62.80%** | 51.20% | 54.00% | 51.60% | **59.20%** |
| LDA | 54.40% | **54.40%** | **61.20%** | 49.60% | 48.80% |
| LIWC + SNA | NAÏVE BAYES | 37.20% | **59.20%** | **65.60%** | 44.40% | 49.60% |
| SVM | 48.40% | 55.20% | 46.80% | 45.20% | 38.40% |
| Logistic Regression | 46.40% | 54.40% | 61.20% | 48.80% | **58.40%** |
| GRADIENT BOOSTING | 56.40% | 54.80% | 57.20% | 53.20% | 57.20% |
| LDA | **56.80%** | 54.40% | 62.80% | **53.60%** | 55.20% |
| SPLICE + SNA | NAÏVE BAYES | 41.20% | 50.40% | 60.00% | 49.20% | 43.20% |
| SVM | 50.80% | 51.20% | 52.40% | 48.40% | 52.00% |
| Logistic Regression | 54.00% | 53.60% | 59.60% | 49.20% | **57.20%** |
| GRADIENT BOOSTING | **56.80%** | 46.80% | **60.80%** | 51.60% | 55.60% |
| LDA | **56.80%** | **55.60%** | 53.20% | **54.00%** | 52.00% |
| LIWC + SPLICE + SNA | NAÏVE BAYES | 41.20% | 50.40% | 60.00% | 49.20% | 43.20% |
| SVM | 50.80% | 51.20% | 52.40% | 48.40% | 52.00% |
| Logistic Regression | 55.60% | 53.60% | 61.60% | 48.80% | 59.20% |
| GRADIENT BOOSTING | **61.20%** | 53.20% | 58.00% | **53.60%** | **60.40%** |
| LDA | 56.40% | **55.60%** | **64.40%** | 50.40% | 50.40% |

Tabel 4.7 diatas merupakan percobaan yang menggunakan *dataset* myPersonality dan 7 fitur berbeda. Percobaan skenario 4 ini menggunakan proses *Oversampling* yaitu menduplikasi data minoritas dalam distribusi *dataset*. Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 65.60% dengan menggunakan algoritma Naïve Bayes. Fitur yang digunakan adalah LIWC dan fitur LIWC+SNA. Hasil akurasi tertinggi kedua adalah 62.80% dengan menggunakan algoritma Gradient Boosting dan fitur LIWC+SPLICE. Penggunaan proses *Oversampling* pada *dataset* myPersonality ini juga tidak dapat meningkatkan hasil akurasi secara signifikan.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion* dengan menggunakan algoritma Naïve Bayes. Walaupun hasil akurasi dengan menggunakan proses *Oversampling* ini menunjukkan penurunan akurasi, *traits* *Conscientiousness* meningkat 0.8% dari akurasi tertinggi sebelumnya untuk *traits* ini.

* + - * 1. **Skenario 5 (Percobaan dengan menggunakan *Features selection* dan dengan *Undersampling*)**

**Tabel 4.8** Hasil Percobaan Skenario 5 dengan *Features selection* dan *Undersampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 48.00% | 48.80% | 57.20% | 46.80% | 36.80% |
| SVM | 55.20% | 52.40% | 54.00% | 53.60% | **54.40%** |
| Logistic Regression | **56.40%** | 50.80% | 54.80% | 53.20% | 52.00% |
| GRADIENT BOOSTING | 42.40% | **53.20%** | 46.00% | **63.20%** | 46.80% |
| LDA | 50.40% | **53.20%** | **59.20%** | 52.80% | 50.00% |
| SPLICE | NAÏVE BAYES | 40.00% | 50.80% | 42.00% | 54.00% | 45.60% |
| SVM | 50.00% | **55.20%** | **52.40%** | **56.80%** | 41.60% |
| Logistic Regression | 45.20% | 53.60% | 50.80% | 51.20% | 48.80% |
| GRADIENT BOOSTING | 46.80% | 51.20% | 47.60% | 52.80% | 54.00% |
| LDA | **51.20%** | 51.60% | 51.60% | 52.00% | **54.40%** |
| SNA | NAÏVE BAYES | 38.00% | **57.60%** | **67.60%** | 48.00% | 50.80% |
| SVM | **44.40%** | 53.60% | 57.60% | 43.60% | 46.00% |
| Logistic Regression | 41.60% | 52.00% | 66.00% | 52.80% | **58.80%** |
| GRADIENT BOOSTING | 43.60% | 50.80% | 62.80% | **55.20%** | 52.80% |
| LDA | 43.60% | 54.40% | 58.40% | 54.80% | 56.80% |
| LIWC + SPLICE | NAÏVE BAYES | 40.00% | 50.40% | 42.00% | 54.00% | 46.40% |
| SVM | 46.80% | 52.40% | 52.80% | **54.80%** | 45.20% |
| Logistic Regression | 44.80% | **52.80%** | 50.00% | 52.00% | 48.80% |
| GRADIENT BOOSTING | 45.20% | 48.80% | 46.00% | 54.00% | **54.80%** |
| LDA | **53.20%** | **52.80%** | **55.20%** | 52.80% | 51.20% |
| LIWC + SNA | NAÏVE BAYES | 38.00% | **57.60%** | **68.00%** | 48.00% | 50.80% |
| SVM | **47.20%** | 49.60% | 51.60% | 43.20% | 45.60% |
| Logistic Regression | 42.80% | 52.00% | 66.00% | 52.80% | **58.00%** |
| GRADIENT BOOSTING | **47.20%** | 55.60% | 59.20% | **56.80%** | 55.20% |
| LDA | 46.00% | 54.40% | 63.60% | 54.80% | 55.60% |
| SPLICE + SNA | NAÏVE BAYES | 41.60% | 51.60% | 53.20% | 52.80% | 42.40% |
| SVM | 44.40% | 51.60% | 58.80% | 52.00% | 50.80% |
| Logistic Regression | 44.00% | **54.80%** | **61.60%** | 52.80% | **57.20%** |
| GRADIENT BOOSTING | **48.40%** | 50.40% | 56.00% | **54.40%** | 56.80% |
| LDA | 48.00% | 52.00% | 53.20% | 53.60% | 55.60% |
| LIWC + SPLICE + SNA | NAÏVE BAYES | 41.60% | 51.60% | 53.20% | 52.80% | 42.00% |
| SVM | 44.40% | 52.00% | 56.80% | 49.20% | 50.00% |
| Logistic Regression | 49.20% | 54.80% | **61.20%** | 51.60% | 54.00% |
| GRADIENT BOOSTING | **52.00%** | 50.80% | 53.60% | 53.20% | **56.80%** |
| LDA | 49.20% | **55.60%** | 55.20% | **54.80%** | 46.80% |

Tabel 4.8 diatas merupakan percobaan yang menggunakan *dataset* myPersonality dan 7 fitur berbeda. Percobaan skenario 5 ini menggunakan gabungan proses *Features selection* dan proses *Undersampling*. Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 68.00% dengan menggunakan algoritma Naïve Bayes. Fitur yang digunakan adalah fitur LIWC+SNA. Hasil akurasi tertinggi kedua adalah 63.20% dengan menggunakan algoritma Gradient Boosting dan fitur LIWC.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion* dengan menggunakan algoritma Naïve Bayes. . Gabungan proses *Features selection* dan *Undersampling* pada skenario ini membuat *traits* *Agreeableness* memiliki akurasi tertinggi untuk *traits* ini dibandingkan dengan skenario-skenario sebelumnya. Sedangkan semua *traits* lain masih memiliki akurasi yang lebih rendah.

Penggabungan dua proses pada skenario ini ternyata tidak dapat meningkatkan hasil akurasi yang baik karena lebih banyak hasil yang mengalami penurunan jika dibandingkan dengan hasil akurasi yang mengalami peningkatan. Hasil skenario ini secara rata-rata sama dengan hasil skenario 2 (Tabel 4.5) yang hanya menggunakan *Features selection* dan skenario 3 (Tabel 4.6) yang hanya menggunakan *Undersampling*.

* + - * 1. **Skenario 6 (Percobaan dengan menggunakan *Features selection* dan dengan *Oversampling*)**

**Tabel 4.9** Hasil Percobaan Skenario 6 dengan *Features selection* dan *Oversampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 54.40% | 48.40% | 56.00% | 47.20% | 37.60% |
| SVM | 55.20% | 52.80% | 54.80% | 52.80% | 51.60% |
| Logistic Regression | 54.80% | 50.00% | 54.00% | 53.60% | 50.80% |
| GRADIENT BOOSTING | **59.20%** | **56.40%** | 51.20% | **59.60%** | **53.20%** |
| LDA | 51.20% | 54.80% | **59.20%** | 55.20% | 50.40% |
| SPLICE | NAÏVE BAYES | 41.20% | 51.20% | 50.80% | **56.00%** | 44.00% |
| SVM | 44.40% | 50.40% | 54.40% | 52.40% | 48.40% |
| Logistic Regression | 51.60% | 53.20% | **55.20%** | 52.00% | 50.80% |
| GRADIENT BOOSTING | **60.40%** | **54.00%** | 54.00% | 52.40% | 55.20% |
| LDA | 58.80% | 53.60% | 54.00% | 51.20% | **57.20%** |
| SNA | NAÏVE BAYES | 37.20% | **59.20%** | 65.60% | 44.00% | 49.60% |
| SVM | 50.40% | 55.20% | 49.60% | 45.20% | 38.40% |
| Logistic Regression | 45.20% | 54.40% | 61.20% | 48.80% | 58.00% |
| GRADIENT BOOSTING | **55.20%** | 50.00% | **66.80%** | **52.00%** | 56.40% |
| LDA | 45.60% | 54.00% | 61.20% | 51.20% | **58.80%** |
| LIWC + SPLICE | NAÏVE BAYES | 41.60% | 52.00% | 53.20% | **56.00%** | 43.60% |
| SVM | 53.60% | 54.00% | 54.40% | 48.80% | 44.80% |
| Logistic Regression | 52.40% | 52.80% | 55.20% | 51.20% | 50.80% |
| GRADIENT BOOSTING | **58.80%** | 50.80% | 54.00% | 54.00% | **53.20%** |
| LDA | 53.60% | **56.00%** | **56.00%** | 52.80% | 52.80% |
| LIWC + SNA | NAÏVE BAYES | 37.20% | **59.20%** | **65.60%** | 44.00% | 49.60% |
| SVM | 50.40% | 55.20% | 49.60% | 45.20% | 38.40% |
| Logistic Regression | 45.20% | 54.00% | 61.20% | 49.20% | **58.00%** |
| GRADIENT BOOSTING | **56.40%** | 57.60% | 59.60% | **59.60%** | 57.60% |
| LDA | 55.60% | 52.40% | 62.80% | 54.40% | 52.80% |
| SPLICE + SNA | NAÏVE BAYES | 41.20% | 50.40% | 60.80% | 50.00% | 43.60% |
| SVM | 50.40% | 51.20% | 54.40% | 49.20% | 49.60% |
| Logistic Regression | 56.80% | **54.80%** | **61.20%** | 48.80% | **58.40%** |
| GRADIENT BOOSTING | **61.20%** | 48.00% | 59.20% | **54.80%** | 56.00% |
| LDA | 56.00% | 54.40% | 58.00% | 52.40% | 57.60% |
| LIWC + SPLICE + SNA | NAÏVE BAYES | 41.20% | 50.40% | **60.80%** | 50.00% | 43.20% |
| SVM | 50.40% | 51.20% | 54.40% | 49.20% | 52.00% |
| Logistic Regression | 56.40% | 55.20% | **60.80%** | 46.40% | 59.20% |
| GRADIENT BOOSTING | **58.00%** | 51.60% | 58.80% | 52.00% | **60.40%** |
| LDA | 56.00% | **56.40%** | 58.00% | **52.80%** | 50.40% |

Tabel 4.9 diatas merupakan percobaan yang menggunakan *dataset* myPersonality dan 7 fitur berbeda. Percobaan skenario 6 ini menggunakan gabungan proses *Features selection* dan proses *Oversampling*. Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 66.80% dengan menggunakan algoritma Gradient Boosting dan fitur SNA. Hasil akurasi tertinggi kedua adalah 61.20% dengan menggunakan algoritma Gradient Boosting dan fitur SPLICE+SNA.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion* dengan menggunakan algoritma Gradient Boosting dan fitur SNA.

Penggabungan dua proses pada skenario ini ternyata tidak dapat meningkatkan hasil akurasi menjadi lebih baik. Hasil skenario ini juga mengalami penurunan secara rata-rata jika dibandingkan dengan skenario 2 (Tabel 4.5) yang hanya menggunakan *Features selection* atau skenario 4 (Tabel 4.7) yang hanya menggunakan *Oversampling*.

* + - * 1. **Kesimpulan Hasil Implementasi *Machine learning* pada *Dataset* my*Personality***

Berikut adalah tabel hasil kesimpulan dari semua skenario percobaan yang menggunakan *dataset* my*Personality*.

**Tabel 4.10** Kesimpulan hasil implementasi *Machine learning* pada *Dataset* my*Personality*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Traits*** | **Accuracy** | ***Algorithm*** | **Skenario** | ***Features*** |
| ***Openness*** | 70.4 | SVM | Skenario 2 | LIWC |
| Logistic Regression | Skenario 1 Skenario 2 |
| ***Conscientiousness*** | 59.2 | Naïve Bayes | Skenario 4 Skenario 6 | LIWC |
| Skenario 4 Skenario 6 | LIWC+SNA |
| ***Extraversion*** | 68.8 | Naïve Bayes | Skenario 1 Skenario 2 | LIWC |
| Skenario 1 Skenario 2 | LIWC+SNA |
| ***Agreeableness*** | 63.2 | Gradient Boosting | Skenario 5 | LIWC |
| ***Neuroticism*** | 60.8 | Linear Discriminant Analysis | Skenario 1 | LIWC |

Dari Tabel kesimpulan 4.10 diatas dapat dilihat bahwa fitur LIWC merupakan fitur yang paling dominan untuk digunakan pada semua *traits* kepribadian *Big Five* *Personality*. Algoritma yang menjadi akurasi tertinggi cukup menyebar dan hanya algoritma Naïve Bayes yang muncul dua kali sebagai algoritma terbaik untuk *traits* *Conscientiousness* dan *traits* *Extraversion*.

Untuk skenario di *dataset* my*Personality* ini, dapat ditarik kesimpulan bahwa ternyata penambahan *Features selection* dan proses *Resampling* tidak secara signifikan membantu hasil akurasi dari percobaan. Tiga *traits* yaitu, *Openness*, *Extraversion*, dan *Neuroticism* masih menggunakan skenario 1 yang tidak menambahkan proses *Features selection* dan proses *Resampling*. *Traits* *Conscientiousness* dan *Agreeableness* juga hanya mengalami peningkatan beberapa persen dari skenario 1.

* + - 1. ***Dataset Manual gathering***

Percobaan dengan skenario 7-12 dilakukan dengan menggunakan *dataset* yang didapatkan secara manual oleh peneliti. Tujuan percobaan dengan *dataset* ini dilakukan secara terpisah untuk mencegah terjadinya ketidakseimbangan hasil penelitian yang menggabungkan *dataset* ini dengan *dataset* my*Personality* (Kosinski, 2015), sehingga peneliti dapat melihat langsung akurasi yang diperoleh oleh *dataset* ini.

Percobaan dengan *dataset* ini juga akan menghilangkan penggunaan fitur SNA yang hanya disediakan oleh *dataset* my*Personality*.

* + - * 1. **Skenario 7 (Percobaan tanpa menggunakan *Features selection* dan tanpa *Resampling*)**

**Tabel 4.11** Hasil Percobaan Skenario 7 tanpa *Features selection* dan tanpa *Resampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 60.67% | 62.67% | 73.33% | 52.00% | 59.33% |
| SVM | 62.67% | 65.33% | 76.00% | 56.67% | **69.33%** |
| Logistic Regression | 64.00% | 57.33% | 74.00% | **58.00%** | 66.67% |
| GRADIENT BOOSTING | **67.33%** | 62.67% | 70.67% | **58.00%** | 66.67% |
| LDA | 58.00% | **67.33%** | **79.33%** | 57.33% | 57.33% |
| SPLICE | NAÏVE BAYES | 42.00% | 44.00% | **70.67%** | 47.33% | 43.33% |
| SVM | 60.67% | 58.00% | 46.67% | 56.67% | 43.33% |
| Logistic Regression | **65.33%** | **66.00%** | 66.00% | 55.33% | 56.00% |
| GRADIENT BOOSTING | 58.67% | 57.33% | 66.67% | **58.67%** | **58.67%** |
| LDA | 58.00% | 61.33% | 61.33% | 54.00% | 58.00% |
| LIWC + SPLICE | NAÏVE BAYES | 42.00% | 44.00% | 69.33% | 46.00% | 44.00% |
| SVM | 58.00% | 55.33% | 47.33% | 58.00% | 54.00% |
| Logistic Regression | 66.00% | **66.00%** | 65.33% | 56.00% | 57.33% |
| GRADIENT BOOSTING | **67.33%** | 61.33% | **72.00%** | 56.00% | **65.33%** |
| LDA | 54.00% | 47.33% | 48.67% | **62.67%** | 56.00% |

Tabel 4.11 diatas merupakan percobaan yang menggunakan *dataset* *Manual gathering* dan 3 fitur berbeda. Evaluasi testing dilakukan dengan 10-fold cross validation. Percobaan skenario 7 ini tidak menggunakan proses *Features selection* dan juga proses Resampling

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 79.33% dengan menggunakan algoritma Linear Discrimant Analysis dan fitur LIWC. Hasil akurasi tertinggi kedua adalah 69.33% dengan menggunakan algoritma Support Vector Machine dan fitur LIWC. Dari hasil tersebut dapat dilihat bahwa fitur linguistik LIWC merupakan fitur yang tepat untuk diaplikasikan pada sistem prediksi karena berhasil mencapai akurasi tertinggi setelah diaplikasikan pada *dataset* myPersonality dan *dataset* *Manual gathering*.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion* dengan fitur LIWC. Hasil *traits* *Extraversion* juga merupakan yang tertinggi di fitur lainnya yaitu fitur SPLICE dan gabungan LIWC+SPLICE.

* + - * 1. **Skenario 8 (Percobaan dengan menggunakan *Features selection*)**

**Tabel 4.12** Hasil Percobaan Skenario 8 dengan *Features selection*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 52.67% | 54.67% | 72.67% | 49.33% | **70.00%** |
| SVM | **64.67%** | 57.33% | 74.67% | 53.33% | 66.67% |
| Logistic Regression | **64.67%** | **58.00%** | 74.67% | **54.00%** | 66.67% |
| GRADIENT BOOSTING | 51.33% | 48.67% | 76.00% | 52.00% | 61.33% |
| LDA | 60.00% | 57.33% | **78.00%** | **54.00%** | 66.67% |
| SPLICE | NAÏVE BAYES | 41.33% | 44.00% | **70.00%** | 46.67% | 40.67% |
| SVM | 60.67% | 54.00% | 52.00% | 55.33% | 54.67% |
| Logistic Regression | **65.33%** | **66.67%** | 66.00% | **59.33%** | 57.33% |
| GRADIENT BOOSTING | 61.33% | 54.67% | 68.00% | 55.33% | 57.33% |
| LDA | 56.67% | 62.00% | 64.00% | 54.00% | **62.00%** |
| LIWC + SPLICE | NAÏVE BAYES | 41.33% | 44.00% | 70.00% | 46.67% | 40.67% |
| SVM | 60.67% | 54.00% | 58.00% | 55.33% | 50.67% |
| Logistic Regression | **66.67%** | **66.67%** | 66.00% | **59.33%** | 56.67% |
| GRADIENT BOOSTING | 62.00% | 56.67% | 72.67% | 58.00% | **64.67%** |
| LDA | 54.67% | 63.33% | **74.00%** | 53.33% | 64.00% |

Tabel 4.12 diatas merupakan percobaan yang menggunakan *dataset* *Manual gathering* dan 3 fitur berbeda. Percobaan skenario 8 ini hanya menggunakan proses *Features selection* seperti halnya yang dilakukan pada skenario 2 (Tabel 4.5) yaitu melakukan pemilihan terhadap fitur-fitur yang memiliki korelasi tinggi dan dianggap optimal terhadap peningkatan hasil akurasi*.* Evaluasi testing dilakukan dengan 10-fold cross validation.

Dari hasil pada tabel diatas, dapat dilihat bahwa di skenario 8 ini nilai akurasi tertinggi yaitu 78.00% dengan algoritma Linear Discriminant Analysis dan fitur LIWC. Algoritma dan fitur yang digunakan sama dengan skenario 1, namun akurasi mengalami penurunan sebesar 1.33% dari skenario sebelumnya. Hasil akurasi tertinggi kedua dengan nilai 70.00% dan menggunakan algoritma Naïve Bayes dan fitur LIWC. Hasil akurasi tertinggi kedua ini menggunakan fitur yang sama namun algoritma yang berbeda dan mengalami peningkatan sebesar 0.67%.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi juga didapatkan oleh *traits* *Extraversion*. Kesimpulan dari penggunaan *Features selection* pada *dataset* ini tidak jauh berbeda dengan pada *dataset* myPersonality (Skenario 2). *Features selection* tidak berhasil meningkatkan hasil akurasi dan masih kalah dibandingkan hasil pada skenario 7. Penggunaan *Features selection* hanya unggul sedikit pada *traits* *Neuroticism* dengan akurasi 70.00%

* + - * 1. **Skenario 9 (Percobaan dengan menggunakan *Undersampling*)**

**Tabel 4.13** Hasil Percobaan Skenario 9 dengan *Undersampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 52.67% | 58.00% | **67.33%** | 50.67% | 46.00% |
| SVM | 55.33% | 58.67% | 56.67% | 56.00% | **59.33%** |
| Logistic Regression | **62.00%** | 54.00% | 50.00% | 57.33% | 52.67% |
| GRADIENT BOOSTING | 53.33% | **62.00%** | 52.67% | 55.33% | 50.67% |
| LDA | 54.67% | 58.67% | 58.00% | **60.67%** | 52.00% |
| SPLICE | NAÏVE BAYES | 41.33% | 44.67% | **60.00%** | 47.33% | 40.67% |
| SVM | 46.00% | 51.33% | 56.00% | 56.00% | 53.33% |
| Logistic Regression | **62.67%** | **62.00%** | 59.33% | **58.00%** | 52.00% |
| GRADIENT BOOSTING | 56.00% | 57.33% | 53.33% | 56.00% | **58.00%** |
| LDA | 52.00% | 57.33% | 51.33% | 50.67% | 52.67% |
| LIWC + SPLICE | NAÏVE BAYES | 40.67% | 44.67% | **62.67%** | 46.00% | 42.00% |
| SVM | 51.33% | 53.33% | 56.00% | 51.33% | 51.33% |
| Logistic Regression | **61.33%** | 62.00% | 58.00% | 57.33% | 53.33% |
| GRADIENT BOOSTING | 58.67% | **66.00%** | 54.67% | **59.33%** | **59.33%** |
| LDA | 50.00% | 64.67% | 62.00% | 46.00% | 57.33% |

Tabel 4.13 diatas merupakan percobaan yang menggunakan *dataset* *Manual gathering* dan 3 fitur berbeda. Percobaan skenario 9 ini menggunakan proses *Undersampling* seperti halnya yang dilakukan pada skenario 3 (Tabel 4.6) yaitu menambahkan data minoritas dalam distribusi *dataset* untuk menyeimbangkan distribusi data. Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 67.33% dengan menggunakan algoritma Naïve Bayes dan fitur LIWC. Hasil akurasi tertinggi kedua adalah 66.00% dengan menggunakan algoritma Gradient Boosting dan fitur gabungan LIWC+SPLICE. Hasil skenario ini menunjukkan penggunaan *Undersampling* pada *dataset* ini juga tidak berhasil meningkatkan hasil akurasi, sebaliknya hasil akurasi mengalami penurunan dengan penggunaan *Undersampling* pada *dataset* *Manual gathering* ini.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi masih didapatkan oleh *traits* *Extraversion* dengan menggunakan algoritma Naïve Bayes (67.33%). *Traits* *Conscientiousness* menjadi *traits* kedua tertinggi setelahnya dengan akurasi 66.00%.

* + - * 1. **Skenario 10 (Percobaan dengan menggunakan *Oversampling*)**

**Tabel 4.14** Hasil Percobaan Skenario 10 dengan *Oversampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 52.67% | 60.67% | 73.33% | 53.33% | 46.00% |
| SVM | 62.00% | 62.00% | 68.00% | 56.00% | **64.00%** |
| Logistic Regression | 61.33% | 60.00% | 67.33% | 55.33% | 60.67% |
| GRADIENT BOOSTING | **67.33%** | 62.67% | 72.00% | 57.33% | 62.67% |
| LDA | 59.33% | **64.00%** | **78.00%** | **59.33%** | 57.33% |
| SPLICE | NAÏVE BAYES | 43.33% | 44.00% | 58.00% | 47.33% | 42.67% |
| SVM | 55.33% | 60.67% | 52.67% | 57.33% | 44.67% |
| Logistic Regression | **62.67%** | **64.67%** | 62.67% | 56.67% | 57.33% |
| GRADIENT BOOSTING | 62.00% | 60.67% | **65.33%** | **58.00%** | **58.67%** |
| LDA | 56.00% | 63.33% | 54.67% | 52.00% | 54.67% |
| LIWC + SPLICE | NAÏVE BAYES | 42.67% | 43.33% | 63.33% | 46.67% | 40.67% |
| SVM | 60.67% | 60.67% | 60.00% | 56.00% | 44.67% |
| Logistic Regression | 63.33% | 66.00% | 61.33% | 56.00% | 58.00% |
| GRADIENT BOOSTING | **64.67%** | **66.67%** | **73.33%** | 55.33% | **67.33%** |
| LDA | 52.67% | 46.67% | 48.67% | **62.67%** | 56.67% |

Tabel 4.14 diatas merupakan percobaan yang menggunakan *dataset* *Manual gathering* dan 3 fitur berbeda. Percobaan skenario 10 ini menggunakan proses *Oversampling* seperti halnya yang dilakukan pada skenario 4 (Tabel 4.7) yaitu mengurangi data mayoritas dalam distribusi *dataset* untuk menyeimbangkan distribusi data. Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 78.00% dengan menggunakan algoritma Linear Discriminant Analysis dan fitur LIWC. Hasil akurasi tertinggi kedua adalah 67.33% dengan menggunakan algoritma Gradient Boosting dan fitur LIWC serta fitur gabungan LIWC+SPLICE. Penggunaan *Oversampling* pada *dataset* ini juga tidak meningkatkan hasil akurasi, namun masih lebih baik jika dibandingkan dengan penggunaan *Undersampling* yang menurunkan hasil akurasi.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi masih didapatkan oleh *traits* *Extraversion* dengan menggunakan algoritma Linear Discriminant Analysis (78.00%). *Traits* *Conscientiousness* dan *Neuroticism* menjadi *traits* kedua tertinggi setelahnya dengan akurasi 67.33% menggunakan Gradient Boosting.

* + - * 1. **Skenario 11 (Percobaan dengan menggunakan *Features selection* dan dengan *Undersampling*)**

Tabel 4.15 dibawah merupakan percobaan yang menggunakan *dataset* *Manual gathering* dan 3 fitur berbeda. Percobaan skenario 11 ini menggabungkan proses *Features selection* dan proses *Undersampling* seperti halnya yang dilakukan pada skenario 5 (Tabel 4.8). Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 72.67% dengan menggunakan algoritma Support Vector Machine dan fitur LIWC. Hasil akurasi tertinggi kedua adalah 64.67% dengan menggunakan algoritma Linear Discriminant Analysis dan fitur gabungan LIWC+SPLICE. Penggunaan proses gabungan *Features selection* dan *Undersampling* ternyata juga tidak berhasil meningkatkan hasil akurasi seperti yang terjadi pada skenario 5 (Tabel 4.8).

Hasil akurasi proses gabungan ini lebih baik dibandingkan dengan skenario 9 yang hanya menggunakan proses *Undersampling*. Namun, jika dibandingkan dengan skenario 8 yang hanya menggunakan *Features selection*, hasil akurasi pada skenario ini jauh lebih rendah. Kesimpulan yang dapat diambil adalah penggunaan *Undersampling* tidak cocok untuk digunakan pada *dataset* ini, setelah dua kali percobaan yang menggunakan proses *Undersampling* menurunkan hasil akurasi.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion*. *Traits* *Conscientiousness* menjadi *traits* kedua tertinggi setelahnya dengan akurasi 64.67% dengan menggunakan algoritma Linear Discriminant Analysis.

**Tabel 4.15** Hasil Percobaan Skenario 11 dengan *Features selection* dan *Undersampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 43.33% | 52.67% | 51.33% | 52.67% | 60.00% |
| SVM | 47.33% | 50.67% | **72.67%** | 55.33% | **64.00%** |
| Logistic Regression | **48.67%** | 52.67% | 69.33% | 53.33% | **64.00%** |
| GRADIENT BOOSTING | 46.67% | **53.33%** | 64.00% | 52.67% | 55.33% |
| LDA | 43.33% | 49.33% | 68.00% | **54.00%** | **64.00%** |
| SPLICE | NAÏVE BAYES | 41.33% | 43.33% | **60.00%** | 46.00% | 40.67% |
| SVM | 54.67% | 54.00% | 50.67% | **60.67%** | 46.00% |
| Logistic Regression | **62.00%** | **62.67%** | 52.00% | 58.67% | 52.67% |
| GRADIENT BOOSTING | 58.00% | 61.33% | 51.33% | 52.00% | **53.33%** |
| LDA | 56.67% | 60.00% | 48.00% | 50.00% | **55.33%** |
| LIWC + SPLICE | NAÏVE BAYES | 41.33% | 43.33% | 60.67% | 46.00% | 40.67% |
| SVM | 54.00% | 54.00% | 56.00% | 57.33% | 48.67% |
| Logistic Regression | **61.33%** | 62.67% | 50.67% | 58.67% | 53.33% |
| GRADIENT BOOSTING | 52.67% | 56.00% | 59.33% | **60.67%** | 62.00% |
| LDA | 54.67% | **64.67%** | **64.00%** | 55.33% | **64.00%** |

* + - * 1. **Skenario 12 (Percobaan dengan menggunakan *Features selection* dan dengan *Oversampling*)**

**Tabel 4.16** Hasil Percobaan Skenario 12 dengan *Features selection* dan *Oversampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 45.33% | **54.00%** | 66.00% | 50.00% | 62.67% |
| SVM | 47.33% | 48.00% | 68.67% | 56.67% | 62.00% |
| Logistic Regression | 47.33% | 49.33% | 66.00% | **59.33%** | 62.67% |
| GRADIENT BOOSTING | **53.33%** | 48.67% | 70.00% | 52.67% | 59.33% |
| LDA | 42.67% | 51.33% | **70.67%** | 52.00% | **63.33%** |
| SPLICE | NAÏVE BAYES | 42.67% | 42.67% | 57.33% | 46.67% | 40.00% |
| SVM | 54.00% | 58.00% | 53.33% | 54.00% | 41.33% |
| Logistic Regression | **63.33%** | **66.67%** | 59.33% | **59.33%** | 56.67% |
| GRADIENT BOOSTING | 60.67% | 62.00% | **66.67%** | 54.67% | **60.67%** |
| LDA | 54.67% | 62.00% | 53.33% | 54.67% | 56.00% |
| LIWC + SPLICE | NAÏVE BAYES | 43.33% | 42.67% | 58.00% | 47.33% | 40.00% |
| SVM | 50.00% | 60.00% | 56.67% | 56.00% | 41.33% |
| Logistic Regression | **66.00%** | **66.00%** | 59.33% | 58.00% | 57.33% |
| GRADIENT BOOSTING | **66.00%** | 58.67% | **72.00%** | **60.67%** | 60.00% |
| LDA | 52.00% | 64.00% | 70.67% | 56.67% | **63.33%** |

Tabel 4.16 diatas merupakan percobaan yang menggunakan *dataset* *Manual gathering* dan 3 fitur berbeda. Percobaan skenario 12 ini menggunakan proses gabungan *Features selection* dan proses *Oversampling* seperti halnya yang dilakukan pada skenario 6 (Tabel 4.9). Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 72.00% dengan menggunakan algoritma Gradient Boosting dan fitur gabungan LIWC+SPLICE. Hasil akurasi tertinggi kedua adalah 66.67% dengan menggunakan algoritma Logistic Regression dan fitur SPLICE. Hasil yang ditunjukkan oleh skenario ini menunjukkan ciri-ciri yang sama dengan skenario 6 (Tabel 4.9) yang juga menggunakan proses gabungan *Features selection* dan *Oversampling*. Akurasi yang tertinggi sama-sama menggunakan algoritma Gradient Boosting dan hasil yang ditunjukkan juga menurun jika dibandingkan dengan proses yang hanya menggunakan *Features selection* (skenario 8) atau proses yang hanya menggunakan *Oversampling* (skenario 10).

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi masih didapatkan oleh *traits* *Extraversion*. *Traits* *Conscientiousness* menjadi *traits* kedua tertinggi setelahnya dengan akurasi 66.67% menggunakan algoritma Logistic Regression.

* + - * 1. **Kesimpulan Hasil Implementasi *Machine learning* pada *Dataset* *Manual gathering***

**Tabel 4.17** Kesimpulan hasil implementasi *Machine learning* pada

*Dataset* *Manual gathering*

| ***Traits*** | **Accuracy** | ***Algorithm*** | **Skenario** | ***Features*** |
| --- | --- | --- | --- | --- |
| ***Openness*** | 67.33 | Gradient Boosting | Skenario 7 | LIWC+SPLICE |
| Skenario 7 Skenario 10 | LIWC |
| ***Conscientiousness*** | 67.33 | Linear Discriminant Analysis | Skenario 7 | LIWC |
| ***Extraversion*** | 79.33 | Linear Discriminant Analysis | Skenario 7 | LIWC |
| ***Agreeableness*** | 62.67 | Linear Discriminant Analysis | Skenario 7 Skenario 10 | LIWC+SPLICE |
| ***Neuroticism*** | 70 | Naïve Bayes | Skenario 8 | LIWC |

Dari tabel kesimpulan diatas dapat dilihat bahwa ternyata benar bahwa skenario yang tidak menggunakan proses *Features selection* dan proses *Resampling* masih mendominasi akurasi tertinggi seperti pada *dataset* *Manual gathering* ini di mana semua *traits* memiliki akurasi tertinggi jika menggunakan skenario 7 kecuali *Neuroticism* yang mendapatkan akurasi tertinggi dengan skenario 8 yang menggunakan proses *Features selection*. Fitur yang paling mendominasi di *dataset* ini masih merupakan fitur LIWC di mana fitur ini digunakan di semua *traits* kepribadian. Untuk algoritma, Linear Discriminant Analysis mendominasi dengan menjadi akurasi tertinggi bagi 3 *traits* yaitu, *Conscientiousness*, *Extraversion*, dan *Agreeableness*.

Jika dibandingkan dengan hasil percobaan skenario 1-6 yang menggunakan *dataset* my*Personality*, akurasi tertinggi di *dataset* manual ini lebih tinggi untuk *traits* *Conscientiousness* (67.33% dibandingkan dengan 59.2%), *Extraversion* (79.33% dibandingkan dengan 68.8%), dan *Neuroticism* (70% dibandingkan dengan 60.8%). Sedangkan dua *traits* lainnya diungguli oleh my*Personality*, *Openness* (67.33% dibandingkan dengan 70.4%) dan *Agreeableness* (62.67% dibandingkan dengan 63.2%).

*Traits* *Extraversion* menjadi *traits* dengan akurasi tertinggi dari semua skenario dengan 79.33%. *Traits* *Extraversion* juga menjadi *traits* dengan akurasi tertinggi di semua skenario percobaan yang menggunakan *Dataset* *Manual gathering* (skenario 7-12) ini.

* + - 1. ***Dataset* Gabungan**

Percobaan dengan skenario 13-18 dilakukan dengan menggunakan *dataset* gabungan dari *dataset* my*Personality* dan *dataset* *Manual gathering* yang berjumlah 400 *dataset*. Jika sebelumnya percobaan dilakukan untuk masing-masing *dataset* untuk mencegah terjadinya ketidakseimbangan data bila digabungkan, maka untuk percobaan kali ini, peneliti mencoba menggabungkan kedua *dataset* tersebut dan melihat bagaiman hasil akurasinya jika dibandingkan dengan *dataset* sebelum digabungkan.

Peneliti ingin mengetahui apakah dengan jumlah data yang lebih besar dapat menghasilkan proses *learning* dan akurasi yang lebih baik atau sebaliknya karena menggunakan *dataset* gabungan yang berkemungkinan menjadi tidak seimbang dan membuat akurasi hasil percobaan semakin menurun.

* + - * 1. **Skenario 13 (Percobaan tanpa menggunakan *Features selection* dan tanpa *Resampling*)**

**Tabel 4.18** Hasil Percobaan Skenario 13 tanpa *Features selection* dan tanpa *Resampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 61.50% | 55.50% | 62.00% | 52.00% | 49.25% |
| SVM | 67.25% | **58.75%** | 66.50% | 54.50% | **62.50%** |
| Logistic Regression | **68.50%** | 54.25% | 66.25% | 51.00% | **62.50%** |
| GRADIENT BOOSTING | 65.25% | 57.25% | 64.25% | 55.75% | 59.75% |
| LDA | 62.50% | 58.00% | **68.50%** | **56.00%** | 55.00% |
| SPLICE | NAÏVE BAYES | 34.50% | 50.25% | 61.25% | 46.50% | 40.75% |
| SVM | 44.75% | 51.50% | 52.00% | **52.50%** | 48.75% |
| Logistic Regression | 63.00% | 52.25% | **63.25%** | 48.50% | 61.25% |
| GRADIENT BOOSTING | 61.50% | 54.50% | 60.75% | 50.00% | 57.25% |
| LDA | **64.25%** | **56.25%** | 60.00% | 51.25% | **63.75%** |
| LIWC + SPLICE | NAÏVE BAYES | 35.00% | 50.50% | 61.50% | 46.00% | 40.00% |
| SVM | 58.25% | 49.25% | 58.00% | 48.75% | 52.25% |
| Logistic Regression | **63.50%** | 55.00% | 62.50% | 47.75% | **63.75%** |
| GRADIENT BOOSTING | **63.50%** | 56.25% | 62.50% | **56.50%** | 59.50% |
| LDA | 58.00% | **59.25%** | **64.25%** | 54.75% | 59.25% |

Tabel 4.18 diatas merupakan percobaan yang menggunakan *dataset* gabungan dari *dataset* myPersonality dan *dataset* *Manual gathering* serta penggunaan 3 fitur berbeda. Percobaan skenario 13 ini tidak menggunakan proses *Features selection* dan juga proses Resampling. Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 68.50% dengan menggunakan algoritma Linear Discrimant Analysis dan algoritma Logistic Regression dengan sama-sama menggunakan fitur LIWC. Hasil akurasi tertinggi kedua adalah 63.75% dengan menggunakan algoritma Logistic Regression dengan fitur gabungan LIWC+SPLICE dan algoritma Linear Discriminant Analysis (LDA) dengan fitur SPLICE.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion* dan *Openness* dengan menggunakan fitur LIWC. Percobaan yang dilakukan pada *dataset* ini yang lebih besar dibandingkan dengan percobaan dengan *dataset* terpisah sebelumnya (skenario 1 untuk *dataset* myPersonality dengan Tabel 4.4 dan skenario 7 untuk *dataset* *manual data gathering* dengan Tabel 4.11) ternyata tidak mengalami peningkatan hasil akurasi. Hasil pada Tabel 4.18 menunjukkan tidak ada hasil akurasi yang lebih tinggi dibandingkan percobaan dengan *dataset* terpisah sebelumnya.

* + - * 1. **Skenario 14 (Percobaan dengan menggunakan *Features selection*)**

**Tabel 4.19** Hasil Percobaan Skenario 14 dengan *Features selection*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 65.50% | 51.25% | 63.50% | 45.50% | 57.00% |
| SVM | 68.00% | 52.25% | 66.25% | 52.50% | 62.25% |
| Logistic Regression | **68.25%** | 50.25% | **66.50%** | 53.75% | **62.75%** |
| GRADIENT BOOSTING | 62.00% | 52.50% | 59.50% | 54.75% | 56.25% |
| LDA | 68.00% | **54.25%** | 66.25% | **55.25%** | 61.50% |
| SPLICE | NAÏVE BAYES | 34.25% | 51.00% | 60.75% | 46.25% | 39.50% |
| SVM | 58.25% | 54.75% | 54.75% | 50.25% | 49.00% |
| Logistic Regression | 63.00% | 54.25% | 64.50% | 49.25% | 62.00% |
| GRADIENT BOOSTING | 62.25% | 54.25% | 58.50% | **51.75%** | 59.50% |
| LDA | **63.50%** | **58.50%** | **61.50%** | 51.50% | **65.25%** |
| LIWC + SPLICE | NAÏVE BAYES | 34.00% | 51.00% | 61.25% | 46.00% | 39.25% |
| SVM | 53.50% | 51.50% | 50.50% | **53.50%** | 48.25% |
| Logistic Regression | **63.25%** | 55.00% | **64.25%** | 49.50% | 62.00% |
| GRADIENT BOOSTING | 63.00% | 55.75% | 57.00% | 53.25% | 57.25% |
| LDA | 62.50% | **58.75%** | 63.00% | 52.00% | **65.00%** |

Tabel 4.19 diatas merupakan percobaan yang menggunakan *dataset* gabungan dari *dataset* myPersonality dan *dataset* *Manual gathering* serta penggunaan 3 fitur berbeda. Percobaan skenario 14 ini hanya menggunakan proses *Features selection* seperti halnya yang dilakukan pada skenario 2 (Tabel 4.5) dan skenario 8 (Tabel 4.12) yaitu melakukan pemilihan terhadap fitur-fitur yang memiliki korelasi tinggi dan dianggap optimal terhadap peningkatan hasil akurasi*.* Evaluasi testing dilakukan dengan 10-fold cross validation.

Dari hasil pada tabel diatas, dapat dilihat bahwa di skenario 14 ini nilai akurasi tertinggi yaitu 68.25% dengan algoritma Logistic Regression dan fitur LIWC. Hasil akurasi tertinggi kedua memiliki nilai 66.50% juga menggunakan algoritma Logistic Regression dan fitur LIWC.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Openness*. Percobaan pada skenario ini menurun dibandingkan dengan hasil pada skenario 13 (Tabel 4.18) yang tidak menggunakan proses apapun, kecuali pada *traits* *Neuroticism* yang mengalami peningkatan sebesar 1.50% dari skenario sebelumnya menjadi 65.25%.

Setelah melakukan proses *Features selection* pada 3 *dataset* berbeda (skenario 2 untuk myPersonality dengan Tabel 4.5, skenario 8 untuk *dataset* *manual data gathering* dengan Tabel 4.12, dan skenario 14 untuk *dataset* gabungan dengan Tabel 4.19), didapatkan kesimpulan bahwa penggunaan *Features selection* tidak meningkatkan akurasi secara signifikan, namun berhasil meningkatkan akurasi tertinggi di salah satu *traits* kepribadian. Peningkatan di *traits* *Agreeableness* untuk *dataset* myPersonality, peningkatan di *traits* *Neuroticism* untuk *dataset* *manual gathering*, dan peningkatan di *traits* *Neuroticism* untuk *dataset* gabungan ini. *Traits* lain hanya mengalami sedikit penurunan akurasi atau memiliki akurasi yang sama dengan skenario yang tidak menggunakan proses apapun.

* + - * 1. **Skenario 15 (Percobaan dengan menggunakan *Undersampling*)**

**Tabel 4.20** Hasil Percobaan Skenario 15 dengan *Undersampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | 39.75% | 55.50% | 49.25% | 54.25% | 39.50% |
| SVM | 42.50% | 58.75% | 52.25% | 52.75% | **52.50%** |
| Logistic Regression | 46.25% | 58.50% | 50.75% | 51.50% | 50.00% |
| GRADIENT BOOSTING | **49.00%** | 55.75% | 52.00% | **56.75%** | 50.50% |
| LDA | 46.00% | **59.50%** | **59.75%** | 56.00% | 49.50% |
| SPLICE | NAÏVE BAYES | 36.75% | 50.50% | **59.75%** | 46.75% | 41.00% |
| SVM | 47.25% | 51.25% | 45.00% | 49.50% | 53.50% |
| Logistic Regression | 48.25% | 52.00% | 52.50% | 50.75% | 51.50% |
| GRADIENT BOOSTING | 47.50% | **56.00%** | 47.25% | **52.00%** | 50.75% |
| LDA | **49.50%** | 55.25% | 49.25% | 51.75% | **54.75%** |
| LIWC + SPLICE | NAÏVE BAYES | 37.75% | 50.50% | **61.00%** | 46.75% | 41.00% |
| SVM | 49.00% | 50.50% | 48.25% | 51.25% | 51.00% |
| Logistic Regression | 47.50% | 54.25% | 52.25% | 50.00% | 53.75% |
| GRADIENT BOOSTING | 48.00% | **58.25%** | 56.25% | **56.00%** | 53.75% |
| LDA | **54.50%** | 57.75% | 60.00% | 53.75% | **56.00%** |

Tabel 4.20 diatas merupakan percobaan yang menggunakan *dataset* gabungan dari *dataset* myPersonality dan *dataset* *Manual gathering* serta penggunaan 3 fitur berbeda. Percobaan skenario 15 ini hanya menggunakan proses *Undersampling* seperti halnya yang dilakukan pada skenario 3 (Tabel 4.6) dan skenario 8 (Tabel 4.13) *.* Evaluasi testing dilakukan dengan 10-fold cross validation.

Dari hasil pada tabel diatas, dapat dilihat bahwa di skenario 15 ini nilai akurasi tertinggi yaitu 61.00% dengan algoritma Naïve Bayes dengan fitur gabungan LIWC+SPLICE. Hasil akurasi tertinggi kedua memiliki nilai 59.50% menggunakan algoritma Linear Discriminant Analysis dan fitur LIWC.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion*. Setelah melakukan proses *Undersampling* pada 3 *dataset* berbeda (skenario 3 untuk myPersonality dengan Tabel 4.6, skenario 9 untuk *dataset* *manual data gathering* dengan Tabel 4.13, dan skenario 15 untuk *dataset* gabungan dengan Tabel 4.20), didapatkan kesimpulan bahwa penggunaan *Undersampling* bekerja lebih baik pada *dataset* ini jika dibandingkan dengan pada *dataset* terpisah yang jumlah datanya lebih sedikit. Penggunaan *Undersampling* pada *dataset* sebelumnya memiliki akurasi yang jauh di bawah percobaan tanpa menggunakan proses apapun, sedangkan percobaan pada *dataset* gabungan ini memiliki akurasi yang dapat dikatakan cukup seimbang.

* + - * 1. **Skenario 16 (Percobaan dengan menggunakan *Oversampling*)**

**Tabel 4.21** Hasil Percobaan Skenario 16 dengan *Oversampling*

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
|
| LIWC | NAÏVE BAYES | 50.25% | 55.00% | 55.00% | 53.25% | 41.00% |
| SVM | **57.25%** | **60.75%** | 58.50% | 52.75% | 54.00% |
| Logistic Regression | 56.50% | 57.50% | 53.50% | 47.75% | **56.50%** |
| GRADIENT BOOSTING | 57.00% | 59.25% | **65.25%** | **58.25%** | 54.25% |
| LDA | 57.00% | 59.25% | 64.75% | 56.50% | 53.00% |
| SPLICE | NAÏVE BAYES | 37.00% | 50.50% | **60.50%** | 46.25% | 39.50% |
| SVM | **61.25%** | 50.50% | 49.50% | 51.00% | 55.50% |
| Logistic Regression | 54.25% | 52.00% | 53.25% | 49.00% | 51.50% |
| GRADIENT BOOSTING | 60.00% | 54.75% | 56.50% | 48.50% | 56.00% |
| LDA | 54.25% | **55.25%** | 52.00% | **52.25%** | **56.75%** |
| LIWC + SPLICE | NAÏVE BAYES | 36.75% | 50.75% | 60.75% | 45.00% | 38.75% |
| SVM | 57.25% | 51.25% | 44.00% | **55.00%** | 56.25% |
| Logistic Regression | 55.00% | 54.75% | 56.50% | 49.75% | 52.25% |
| GRADIENT BOOSTING | **61.50%** | **59.75%** | **62.25%** | 54.75% | **57.75%** |
| LDA | 54.75% | 58.75% | 59.50% | 53.50% | 56.00% |

Tabel 4.21 diatas merupakan percobaan yang menggunakan *dataset* gabungan dari *dataset* myPersonality dan *dataset* *Manual gathering* serta penggunaan 3 fitur berbeda. Percobaan skenario 16 ini hanya menggunakan proses *Oversampling* seperti halnya yang dilakukan pada skenario 4 (Tabel 4.7) dan skenario 8 (Tabel 4.14) *.* Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 65.25% dengan menggunakan algoritma Gradient Boosting dan fitur LIWC. Hasil akurasi tertinggi kedua adalah 61.50% juga menggunakan algoritma Gradient Boosting tetapi dengan fitur LIWC+SPLICE.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion*. Sedangkan, *traits* tertinggi selanjutnya adalah *traits* *Openness*.

Percobaan skenario 16 yang menggunakan *Oversampling* memperlihatkan hasil yang lebih baik dibandingkan dengan proses *Undersampling*. Ini terbukti dari hasil pada Tabel 4.21 di mana semua *traits* mengalami peningkatan akurasi dibandingkan dengan Tabel 4.20.

Setelah melakukan proses *Oversampling* pada 3 *dataset* berbeda (skenario 4 untuk myPersonality dengan Tabel 4.7, skenario 10 untuk *dataset* *manual data gathering* dengan Tabel 4.14, dan skenario 16 untuk *dataset* gabungan dengan Tabel 4.21), didapatkan kesimpulan bahwa penggunaan *Oversampling* menghasilkan akurasi yang cukup sebanding dengan percobaan tanpa proses apapun. Namun, penggunaan *Oversampling* yang dapat dikatakan paling bagus adalah pada skenario 10 yaitu *dataset* *manual gathering* yang memiliki *dataset* paling kecil. Ini membuktikan bahwa penambahan data sekaligus proses penyeimbangan data pada *dataset* yang kecil membantu meningkatkan hasil akurasi di Implementasi machine Learning.

* + - * 1. **Skenario 17 (Percobaan dengan menggunakan *Features selection* dan dengan *Undersampling*)**

**Tabel 4.22** Hasil Percobaan Skenario 17 dengan *Features selection* dan *Undersampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | **55.75%** | 52.00% | **63.50%** | 45.75% | 37.50% |
| SVM | 51.75% | 50.75% | 58.25% | 52.75% | 51.25% |
| Logistic Regression | 51.75% | 49.75% | 57.00% | 54.50% | 52.00% |
| GRADIENT BOOSTING | 49.00% | 51.50% | 51.00% | 53.00% | **52.50%** |
| LDA | 49.50% | **53.50%** | 57.50% | **57.00%** | 51.75% |
| SPLICE | NAÏVE BAYES | 36.75% | 51.25% | **59.25%** | 47.25% | 41.00% |
| SVM | 49.75% | 50.25% | 53.50% | **54.25%** | 54.50% |
| Logistic Regression | 48.50% | 52.75% | 51.50% | 50.25% | 51.25% |
| GRADIENT BOOSTING | 47.25% | **56.25%** | 47.25% | 51.50% | 51.00% |
| LDA | **50.25%** | 56.00% | 48.00% | 50.50% | **55.50%** |
| LIWC + SPLICE | NAÏVE BAYES | 37.00% | 51.00% | **59.25%** | 47.00% | 40.75% |
| SVM | 46.75% | 53.25% | 43.50% | 51.25% | 54.50% |
| Logistic Regression | 47.75% | 53.50% | 51.50% | 50.75% | 51.50% |
| GRADIENT BOOSTING | 48.25% | 57.00% | 50.50% | **53.50%** | 50.75% |
| LDA | **48.50%** | **58.50%** | 50.00% | 52.25% | **58.25%** |

Tabel 4.22 diatas merupakan percobaan yang menggunakan *dataset* gabungan dari *dataset* myPersonality dan *dataset* *Manual gathering* serta penggunaan 3 fitur berbeda. Percobaan skenario 17 ini menggunakan proses gabungan *Features selection* dan *Undersampling* seperti halnya yang dilakukan pada skenario 5 (Tabel 4.8) dan skenario 9 (Tabel 4.15) *.* Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 63.50% . Sedangkan hasil akurasi tertinggi kedua adalah 58.50%. Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion* dengan menggunakan algoritma Naïve Bayes dan fitur LIWC. *Traits* dengan akurasi tertinggi kedua yaitu *traits* *Conscientiousness* dengan menggunakan algoritma Linear Discriminant Analysis dan fitur gabungan LIWC+SPLICE.

Setelah melakukan proses gabungan *Features selection* dan *Undersampling* pada 3 *dataset* berbeda (skenario 5 untuk *dataset* myPersonality dengan Tabel 4.8, skenario 11 untuk *dataset* *manual data gathering* dengan Tabel 4.15, dan skenario 17 untuk *dataset* gabungan dengan Tabel 4.22), didapatkan kesimpulan bahwa penggunaan proses gabungan ini tidak berhasil dalam meningkatkan hasil akurasi secara keseluruhan. Proses gabungan ini memiliki hasil lebih baik jika dibandingkan dengan percobaan yang hanya menggunakan proses *Undersampling*, tetapi lebih rendah jika dibandingkan dengan percobaan yang hanya menggunakan *Features selection*.

* + - * 1. **Skenario 18 (Percobaan dengan menggunakan *Features selection* dan dengan *Oversampling*)**

**Tabel 4.23** Hasil Percobaan Skenario 18 dengan *Features selection* dan *Oversampling*

| ***Features*** | ***Algorithm*** | **O** | **C** | **E** | **A** | **N** |
| --- | --- | --- | --- | --- | --- | --- |
|
|  |
| LIWC | NAÏVE BAYES | **61.75%** | **52.25%** | **64.25%** | 45.50% | 43.25% |
| SVM | 52.75% | 50.50% | 60.00% | 55.50% | 53.75% |
| Logistic Regression | 54.25% | 49.25% | 57.25% | 54.50% | **56.25%** |
| GRADIENT BOOSTING | 56.50% | 51.50% | 53.00% | 55.75% | 51.50% |
| LDA | 53.00% | 52.00% | 59.25% | **58.00%** | 53.25% |
| SPLICE | NAÏVE BAYES | 37.25% | 51.00% | **60.50%** | 46.25% | 39.00% |
| SVM | 51.25% | 54.75% | 52.00% | 49.25% | 53.50% |
| Logistic Regression | 55.00% | 54.25% | 53.00% | **50.75%** | 53.75% |
| GRADIENT BOOSTING | **57.50%** | 55.00% | 54.25% | **50.75%** | 55.00% |
| LDA | 53.00% | **57.50%** | 54.00% | 50.25% | **56.50%** |
| LIWC + SPLICE | NAÏVE BAYES | 37.25% | 51.00% | **60.75%** | 46.00% | 38.50% |
| SVM | 50.25% | 50.25% | 57.25% | 53.25% | 51.75% |
| Logistic Regression | 54.50% | 55.50% | 54.50% | 51.00% | 54.75% |
| GRADIENT BOOSTING | **59.75%** | 54.50% | 57.00% | **53.75%** | 58.50% |
| LDA | 53.50% | **58.50%** | 54.00% | 51.00% | **59.25%** |

Tabel 4.23 diatas merupakan percobaan yang menggunakan *dataset* gabungan dari *dataset* myPersonality dan *dataset* *Manual gathering* serta penggunaan 3 fitur berbeda. Percobaan skenario 18 ini menggunakan proses gabungan *Features selection* dan *Oversampling* seperti halnya yang dilakukan pada skenario 6 (Tabel 4.9) dan skenario 10 (Tabel 4.16) *.* Evaluasi testing dilakukan dengan 10-fold cross validation.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 64.25% . Sedangkan hasil akurasi tertinggi kedua adalah 61.75%. Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits* *Extraversion* dengan menggunakan algoritma Naïve Bayes dan fitur LIWC. *Traits* dengan akurasi tertinggi kedua yaitu *traits* *Openness* juga menggunakan algoritma Naïve Bayes dan fitur LIWC.

Setelah melakukan proses gabungan *Features selection* dan *Oversampling* pada 3 *dataset* berbeda (skenario 6 untuk *dataset* myPersonality dengan Tabel 4.9, skenario 12 untuk *dataset* *manual data gathering* dengan Tabel 4.16, dan skenario 18 untuk *dataset* gabungan dengan Tabel 4.23), didapatkan kesimpulan bahwa penggunaan proses gabungan ini memiliki akurasi yang lebih baik dibandingkan proses gabungan *Features selection* dan *Undersampling*. Meskipun begitu, proses gabungan ini tidak lebih baik daripada percobaan yang hanya menggunakan *Features selection* dan percobaan yang hanya menggunakan *Oversampling*.

* + - * 1. **Kesimpulan hasil implementasi *Machine learning* pada *Dataset* Gabungan**

Dari tabel kesimpulan dibawah, dapat dilihat bahwa ternyata setelah *dataset* digabungkan, tidak terjadi peningkatan akurasi secara signifikan. Hasil yang dapat dilihat pada Tabel 4.24 menunjukkan tidak ada *traits* yang memiliki akurasi lebih tinggi dibandingkan dengan percobaan yang menggunakan *dataset* terpisah. Akurasi tertinggi yang didapatkan adalah *traits* *Openness* dengan akurasi 68.5% menggunakan algoritma Logistic Regression dan *Extraversion* juga dengan akurasi 68.5% menggunakan algotirma Linear Discriminant Analysis.

Proses yang dapat meningkatkan akurasi pada *dataset* ini adalah proses *Oversampling* yang berhasil menajdi akurasi tertinggi untuk *traits* *Conscientiousness* dan *Agreeableness*. Penggunaan proses *Features selection* juga berhasil meningkatkan akurasi untuk *traits* *Neuroticism*. Proses yang tidak terbukti berhasil untuk *dataset* ini adalah *Undersampling* yang secara keseluruhan membuat hasil akurasi mengalami penurunan.

**Tabel 4.24** Kesimpulan hasil implementasi *Machine learning* pada *Dataset* *Gabungan*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ***Traits*** | **Accuracy** | ***Algorithm*** | **Skenario** | ***Features*** |
| ***Openness*** | 68.5 | Logistic Regression | Skenario 13 | LIWC |
| ***Conscientiousness*** | 60.75 | Support Vector Machine | Skenario 16 | LIWC |
| ***Extraversion*** | 68.5 | Linear Discriminant Analysis | Skenario 13 | LIWC |
| ***Agreeableness*** | 58.25 | Gradient Boosting | Skenario 16 | LIWC |
| ***Neuroticism*** | 65.25 | Linear Discriminant Analysis | Skenario 14 | SPLICE |

* + - 1. **Kesimpulan Implementasi *Machine learning***

Setelah melakukan implementasi metode *Machine learning* pada tiga *dataset* berbeda dengan 6 skenario di masing-masing *dataset* tersebut. Peneliti membuat sebuah kesimpulan yang diperoleh dari seluruh percobaan tersebut. Seluruh akurasi tertinggi yang didapatkan dikumpulkan dan digabungkan menjadi 1. Akurasi tertinggi dengan nilai yang sama akan diutamakan dengan yang menggunakan proses skenario paling sedikit yaitu tidak menggunakan proses sama sekali.

Dari Tabel 4.25 di bawah dapat dilihat bahwa fitur LIWC menjadi fitur paling dominan untuk mendapatkan akurasi yang paling tinggi di implementasi *Machine learning* pada penelitian ini. Faktor Skenario yang berupa dua proses yaitu *Features selection* dan *Resampling* yang dicoba untuk diterapkan dalam penelitian kali ini ternyata tidak memiliki pengaruh besar terhadap peningkatan akurasi. Dari semua tabel kesimpulan di masing-masing *dataset* sebelumnya, hasil akurasi tertinggi didominasi oleh skenario yang tidak menggunakan proses apapun, begitu juga dengan tabel diatas di mana tiga *traits* yaitu, *Openness*, *Conscientiousness*, dan *Extraversion* memiliki akurasi paling tinggi di implementasi *Machine learning* tanpa menggunakan *Features selection* dan *Resampling* apapun. Sedangkan untuk *traits* *Agreeableness* menggunakan *Features selection* dan *Undersampling* untuk mendapatkan akurasi 63.2% dan hanya lebih tinggi 0.53% dari hasil yang tidak menggunakan proses apapun. *Traits* *Neuroticism* mendapatkan akurasi tertinggi yaitu 70% dengan menggunakan *Features selection* dan unggul cukup signifikan sebanyak 9.20% dari hasil yang tidak menggunakan proses apapun.

**Tabel 4.25** Kesimpulan hasil implementasi *Machine learning*

| ***Traits*** | **Accuracy** | ***Algorithm*** | **Proses** | ***Features*** |
| --- | --- | --- | --- | --- |
| ***Openness*** | 70.4 | Logistic Regression | Tanpa *Features selection* dan Tanpa *Resampling* | LIWC |
| ***Conscientiousness*** | 67.33 | Linear Discriminant Analysis | Tanpa *Features selection* dan Tanpa *Resampling* | LIWC |
| ***Extraversion*** | 79.33 | Linear Discriminant Analysis | Tanpa *Features selection* dan Tanpa *Resampling* | LIWC |
| ***Agreeableness*** | 63.2 | Gradient Boosting | *Features selection* dan *Undersampling* | LIWC |
| ***Neuroticism*** | 70 | Naïve Bayes | *Features selection* | LIWC |

Untuk kesimpulan algoritma yang digunakan, Linear Discriminant Analysis atau LDA menjadi salah satu algoritma yang paling dominan dengan selalu muncul di setiap tabel kesimpulan masing-masing *dataset*. LDA mendapatkan akurasi tertinggi di implementasi *Machine learning* untuk dua *traits* kepribadian yaitu, *Conscientiousness* dengan akurasi 67.33% dan *Extraversion* dengan akurasi 79.33%. Untuk *traits* *Openness* menggunakan algoritma Logistic Regression dengan hasil akurasi 70.4%, *traits* *Agreeableness* dengan algoritma Gradient Boosting dan mendapatkan hasil akurasi 63.2%, serta *traits* terakhir yaitu *Neuroticism* yang memperoleh 70% dengan algoritma Naïve Bayes. Jadi, masing-masing algoritma ternyata memiliki kelebihan di masing-masing *traits* dan tidak ada satu algoritma tertentu yang berpengaruh sangat besar terhadap hasil akurasi dari semua *traits* penelitian ini.

Dari Gambar 4.1 dapat dilihat grafik yang menunjukkan perbandingan algoritma *machine learning* yang digunakan pada penelitian ini untuk setiap traits kepribadian *Big Five Personality*. Traits Extraversion memiliki rata-rata akurasi paling tinggi dibandingkan *traits* lainnya. Hasil yang didapatkan dari implementasi *machine learning* ini akan dibandingkan dengan hasil dari implementasi *deep learning* untuk menentukan algoritma, fitur dan proses yang akan diterapkan di hasil akhir aplikasi sistem prediksi kepribadian.

**Gambar 4.1** Evaluasi perbandingan hasil implementasi algoritma tradisional *machine learning* untuk setiap *traits* kepribadian

### Implementasi *Deep learning*

* + - 1. **Skenario**

Implementasi *deep learning* pada *testing* kali ini akan menggunakan 4 model arsitektur yaitu MultiLayer Perceptron (MLP), LSTM, CNN dan GRU. Sebagai tambahan dari referensi, implementasi *Deep learning* kali ini akan menggunakan arsitektur tambahan yaitu gabungan dari CNN dan LSTM. Fitur yang digunakan dalam implementasi *Deep learning* hanya dengan metode *Open Vocabulary*. Proses *testing* pada *Deep learning* juga akan diaplikasikan ke masing-masing *dataset* berbeda seperti halnya implementasi *Machine learning* pada Bagian 4.2.3 diatas. Distribusi data yang akan digunakan dapat kembali di lihat di Bagian 4.2.1.

Dari faktor skenario diatas pada Bagian 4.2.2, *deep learning* hanya akan menggunakan satu faktor tersebut yaitu *Resampling* (Bagian 4.2.2.2) yang terbagi menjadi dua proses yaitu *Undersampling* dan *Oversampling*, sehingga dapat dibentuk skenario percobaan untuk *deep learning* sebagai berikut:

* + 1. Penggunaan *Dataset* my*Personality*, percobaan tanpa menggunakan *Resampling*.
    2. Penggunaan *Dataset* my*Personality*, percobaan dengan menggunakan *Undersampling*.
    3. Penggunaan *Dataset* my*Personality*, percobaan dengan menggunakan *Oversampling*.
    4. Penggunaan *Dataset* *Manual gathering*, percobaan tanpa menggunakan *Resampling*.
    5. Penggunaan *Dataset* *Manual gathering*, percobaan dengan menggunakan *Undersampling*.
    6. Penggunaan *Dataset* *Manual gathering*, percobaan dengan menggunakan *Oversampling*.
    7. Penggunaan *Dataset* Gabungan, percobaan tanpa menggunakan *Resampling*.
    8. Penggunaan *Dataset* Gabungan, percobaan dengan menggunakan *Undersampling*.
    9. Penggunaan *Dataset* Gabungan, percobaan dengan menggunakan *Oversampling*.

Keterangan pada tabel percobaan:

|  |  |
| --- | --- |
|  | = Header Tabel percobaan yang menggunakan *dataset* my*Personality* |
|  | = Header Tabel percobaan yang menggunakan *dataset* *Manual gathering* |
|  | = Header Tabel percobaan yang menggunakan *dataset* gabungan |
|  |  |
| O | = *Traits* *Openness* |
| C | = *Traits* *Conscientiousness* |
| E | = *Traits* *Extraversion* |
| A | = *Traits* *Agreeableness* |
| N | = *Traits* *Neuroticism* |
| **Highlight** | = Akurasi tertinggi pada sebuah *Traits* |

* + - 1. ***Dataset* my*Personality***

Percobaan dengan implementasi *Deep learning* skenario 1-3 akan menggunakan *dataset* my*Personality*.

* + - * 1. **Skenario 1 (Percobaan tanpa menggunakan *Resampling*)**

Tabel 4.26 di bawah merupakan percobaan yang menggunakan *dataset* myPersonality. Percobaan skenario 1 ini tidak menggunakan proses *Resampling*.

**Tabel 4.26** Hasil Percobaan Skenario 1 tanpa proses *Resampling*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Arsitektur | O | C | E | A | N |
| MLP | **68.00%** | 58.00% | 56.00% | 42.00% | 62.00% |
| LSTM | **68.00%** | 52.00% | 58.00% | **52.00%** | 58.00% |
| GRU | **68.00%** | **62.00%** | 58.00% | 50.00% | **64.00%** |
| CNN 1D | 62.00% | 46.00% | 56.00% | **52.00%** | 58.00% |
| CNN 1D+LSTM | 64.00% | 46.00% | **60.00%** | 42.00% | 58.00% |

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 68.00% dengan menggunakan arsitektur MLP, LSTM, dan GRU. Hasil akurasi tertinggi kedua adalah 64.00% dengan menggunakan arsitektur GRU.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits Openness* dengan menggunakan arsitektur MLP, LSTM, dan GRU.

* + - * 1. **Skenario 2 (Percobaan dengan menggunakan *Undersampling*)**

**Tabel 4.27** Hasil Percobaan Skenario 2 dengan Proses *Undersampling*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Arsitektur | O | C | E | A | N |
| MLP | **79.31%** | 50.00% | **78.95%** | 56.52% | **79.49%** |
| LSTM | 44.83% | 45.83% | 57.89% | 56.52% | 56.41% |
| GRU | 31.03% | **54.17%** | 44.74% | 65.22% | 48.72% |
| CNN 1D | **79.31%** | 47.92% | 57.89% | **67.39%** | 61.54% |
| CNN 1D+LSTM | 75.86% | 47.92% | 71.05% | 43.48% | 58.97% |

Tabel 4.27 di atas merupakan percobaan yang menggunakan *dataset* myPersonality. Percobaan skenario 2 ini menggunakan proses *Undersampling* yaitu menghilangkan data mayoritas dalam distribusi *dataset*.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 79.49% dengan menggunakan arsitektur MLP. Hasil akurasi tertinggi kedua adalah 79.31% dengan menggunakan arsitektur MLP dan CNN 1D. Hasil skenario ini menunjukkan penggunaan *Undersampling* pada *dataset* myPersonality ini berhasil meningkatkan hasil akurasi secara signifikan.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan pada *traits Neuroticism* dengan menggunakan arsitektur MLP dan CNN 1D. Hasil yang didapatkan pada skenario ini menunjukkan terdapat beberapa *traits* yang mengalami peningkatan akurasi yaitu *traits Openness* dengan 79.31% meningkat 11.31% pada skenario sebelumnya, *traits Extraversion* dengan 78.95% meningkat 18.95% pada skenario sebelumnya, *traits Agreeableness* dengan 67.39% meningkat 15.39% pada skenario sebelumnya, dan *traits Neuroticism* dengan 79.49% meningkat 15.49% pada skenario sebelumnya.

* + - * 1. **Skenario 3 (Percobaan dengan menggunakan *Oversampling*)**

**Tabel 4.28** Hasil Percobaan Skenario 3 dengan Proses *Oversampling*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Arsitektur | O | C | E | A | N |
| MLP | 54.41% | **59.62%** | 53.33% | 50.00% | 38.89% |
| LSTM | 48.53% | 51.92% | 57.14% | **53.85%** | **58.62%** |
| GRU | 52.94% | **59.62%** | 48.21% | 48.08% | 56.90% |
| CNN 1D | 50.00% | 50.00% | **60.94%** | 50.00% | 40.74% |
| CNN 1D+LSTM | **60.29%** | 57.69% | 50.00% | 50.00% | 53.45% |

Tabel 4.28 diatas merupakan percobaan yang menggunakan *dataset* myPersonality. Percobaan skenario 3 ini menggunakan proses *Oversampling* yaitu menduplikasi data minoritas dalam distribusi *dataset*.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 60.94% dengan menggunakan arsitektur CNN 1D. Hasil akurasi tertinggi kedua adalah 60.29% dengan menggunakan arsitektur CNN 1D+LSTM. Hasil skenario ini menunjukkan penggunaan *Oversampling* pada *dataset* myPersonality ini tidak berhasil meningkatkan hasil akurasi secara signifikan.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits Extraversion* dengan menggunakan arsitektur CNN 1D. Hasil yang didapatkan pada skenario ini menunjukkan hanya satu *traits* yang mengalami peningkatan akurasi yaitu *traits Conscientiousness* dengan 59.62% meningkat 5.45% pada skenario sebelumnya.

* + - * 1. **Kesimpulan Hasil Implementasi *Deep learning* pada *Dataset* my*Personality***

**Tabel 4.29** Kesimpulan hasil implementasi *Deep learning* pada *Dataset* my*Personality*

|  |  |  |  |
| --- | --- | --- | --- |
| ***Traits*** | **Accuracy** | **Arsitektur** | **Skenario** |
| ***Openness*** | 79.31 | MLP | Skenario 2 |
| CNN 1D |
| ***Conscientiousness*** | 62 | GRU | Skenario 1 |
| ***Extraversion*** | 78.95 | MLP | Skenario 2 |
| ***Agreeableness*** | 67.39 | CNN 1D | Skenario 2 |
| ***Neuroticism*** | 79.49 | MLP | Skenario 2 |

Dari Tabel kesimpulan 4.29 di atas dapat dilihat bahwa arsitektur dengan akurasi tertinggi pada masing-masing *traits* didominasi oleh MLP pada *traits Openness, Extraversion,* dan *Neuroticism*, CNN 1D pada *traits Openness* dan *Agreeableness*, dan GRU pada *traits Conscientiousness*.

Untuk skenario pada *dataset* myPersonality ini, dapat ditarik kesimpulan bahwa penggunaan proses *Undersampling* ternyata dapat meningkatkan akurasi secara signifikan. Empat *traits* yaitu *Openness, Extraversion, Agreeableness*, dan *Neuroticism* pada skenario 2 yang menggunakan proses *Undersampling* dan hanya *traits Conscientiousness* pada skenario 1 yang tidak menggunakan proses *Resampling*.

* + - 1. ***Dataset Manual gathering***

Percobaan dengan implementasi *Deep learning* skenario 4-6 akan menggunakan *dataset* *Manual gathering*.

* + - * 1. **Skenario 4 (Percobaan tanpa menggunakan *Resampling*)**

**Tabel 4.30** Hasil Percobaan Skenario 4 tanpa Proses *Oversampling*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Arsitektur | O | C | E | A | N |
| MLP | **66.67%** | 56.67% | **80.00%** | 53.33% | 73.33% |
| LSTM | 63.33% | 53.33% | 70.00% | 53.33% | 73.33% |
| GRU | 63.33% | 56.67% | 70.00% | 46.67% | **76.67%** |
| CNN 1D | 50.00% | 60.00% | 73.33% | **63.33%** | 66.67% |
| CNN 1D+LSTM | **66.67%** | **66.67%** | 73.33% | **63.33%** | 63.33% |

Tabel 4.30 di atas merupakan percobaan yang menggunakan *dataset* *Manual gathering*. Percobaan skenario 4 ini tidak menggunakan proses *Resampling*.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 80.00% dengan menggunakan arsitektur MLP. Hasil akurasi tertinggi kedua adalah 76.67% dengan menggunakan arsitektur GRU.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits Extraversion* dengan menggunakan arsitektur MLP.

* + - * 1. **Skenario 5 (Percobaan dengan menggunakan *Undersampling*)**

**Tabel 4.31** Hasil Percobaan Skenario 5 dengan Proses *Undersampling*

| Arsitektur | O | C | E | A | N |
| --- | --- | --- | --- | --- | --- |
| MLP | 66.67% | 64.00% | **93.33%** | **70.37%** | 75.00% |
| LSTM | 61.90% | 64.00% | 66.67% | 66.67% | 75.00% |
| GRU | 61.90% | 56.00% | 73.33% | 44.44% | 65.00% |
| CNN 1D | **76.19%** | **68.00%** | 86.67% | 55.56% | 75.00% |
| CNN 1D+LSTM | 66.67% | 48.00% | **93.33%** | 62.96% | **80.00%** |

Tabel 4.31 di atas merupakan percobaan yang menggunakan *dataset* *Manual gathering*. Percobaan skenario 5 ini menggunakan proses *Undersampling*.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 93.33% dengan menggunakan arsitektur MLP dan CNN 1D+LSTM. Hasil skenario ini menunjukkan penggunaan *Undersampling* pada *dataset* *Manual gathering* ini berhasil meningkatkan hasil akurasi secara signifikan.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan pada *traits Extraversion* dengan menggunakan arsitektur MLP dan CNN 1D+LSTM. Hasil yang didapatkan pada skenario ini menunjukkan terdapat beberapa *traits* yang mengalami peningkatan akurasi yaitu *traits Openness* dengan 76.19% meningkat 9.52% pada skenario sebelumnya, *traits Conscientiousness* dengan 68.00% meningkat 1.33% pada skenario sebelumnya, *traits Extraversion* dengan 93.33% meningkat 13.33% pada skenario sebelumnya, *traits Agreeableness* dengan 70.37% meningkat 7.04% pada skenario sebelumnya, dan *traits Neuroticism* dengan 80.00% meningkat 3,33% pada skenario sebelumnya. *Dataset* *Manual gathering* yang dilakukan proses *Undersampling* berhasil meningkatkan akurasi pada kelima *traits* jika dibandingkan dengan *dataset* *Manual gathering* tanpa dilakukan proses *Resampling*.

* + - * 1. **Skenario 6 (Percobaan dengan menggunakan *Oversampling*)**

**Tabel 4.32** Hasil Percobaan Skenario 6 dengan Proses *Oversampling*

| Arsitektur | O | C | E | A | N |
| --- | --- | --- | --- | --- | --- |
| MLP | 55.00% | 50.00% | 64.29% | 52.94% | 45.45% |
| LSTM | **67.50%** | 52.94% | 59.52% | 59.38% | **59.09%** |
| GRU | 47.50% | **61.76%** | 59.52% | 59.38% | 56.82% |
| CNN 1D | 60.00% | 59.38% | **83.33%** | **61.76%** | 43.18% |
| CNN 1D+LSTM | **67.50%** | **61.76%** | 73.81% | 59.38% | 56.82% |

Tabel 4.32 di atas merupakan percobaan yang menggunakan *dataset* *Manual gathering*. Percobaan skenario 6 ini menggunakan proses *Oversampling*.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 83.33% dengan menggunakan arsitektur CNN 1D. Hasil akurasi tertinggi kedua adalah 67.50% dengan menggunakan arsitektur LSTM dan CNN 1D+LSTM. Hasil yang didapatkan pada skenario ini menunjukkan penggunaan *Oversampling* pada *dataset* *Manual gathering* ini tidak berhasil meningkatkan hasil akurasi.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits Extraversion* dengan menggunakan arsitektur CNN 1D. Hasil yang didapatkan pada skenario ini menunjukkan bahwa tidak ada *traits* yang mengalami peningkatan akurasi dari skenario sebelumnya.

* + - * 1. **Kesimpulan Hasil Implementasi *Deep learning* pada *Dataset* *Manual gathering***

**Tabel 4.33** Kesimpulan hasil implementasi *Deep learning* pada *Dataset* *Manual gathering*

| ***Traits*** | **Accuracy** | **Arsitektur** | **Skenario** |
| --- | --- | --- | --- |
| ***Openness*** | 76.19% | CNN 1D | Skenario 5 |
| ***Conscientiousness*** | 68.00% | CNN 1D | Skenario 5 |
| ***Extraversion*** | 93.33% | MLP | Skenario 5 |
| CNN 1D + LSTM |
| ***Agreeableness*** | 70.37% | MLP | Skenario 5 |
| ***Neuroticism*** | 80.00% | CNN 1D + LSTM | Skenario 5 |

Hasil pada Tabel 4.33 diatas semakin memperkuat kesimpulan jika penggunaan proses *Undersampling* sangat efektif pada metode *Deep learning* ini. Semua akurasi tertinggi yang di peroleh di percobaan dengan *dataset* manual ini didapatkan dari skenario 5 yang menerapkan proses *Undersampling*. CNN 1D menjadi arsitektur yang mendominasi karena menjadi arsitektur yang berperan mendapatkan akurasi tertinggi untuk 4 *traits* kepribadian, meskipun digabungkan dengan LSTM untuk *traits* *Extraversion* dan *Neuroticism*. Arsitektur MLP juga mendapatkan akurasi tertinggi untuk *Extraversion* dan *Neuroticism*.

* + - 1. ***Dataset* Gabungan**

Percobaan dengan implementasi *Deep learning* skenario 7-9 akan menggunakan *dataset* gabungan.

* + - * 1. **Skenario 7 (Percobaan tanpa menggunakan *Resampling*)**

**Tabel 4.34** Hasil Percobaan Skenario 7 tanpa Proses *Resampling*

| Arsitektur | O | C | E | A | N |
| --- | --- | --- | --- | --- | --- |
| MLP | 62.50% | 52.50% | 65.00% | **57.50%** | 56.25% |
| LSTM | 58.75% | 52.50% | **66.25%** | 48.75% | 61.25% |
| GRU | 65.00% | 55.00% | 61.25% | 46.25% | **63.75%** |
| CNN 1D | **66.25%** | **58.75%** | 65.00% | **57.50%** | 52.50% |
| CNN 1D+LSTM | 61.25% | 52.50% | 62.50% | 55.00% | 41.25% |

Tabel 4.34 di atas merupakan percobaan yang menggunakan *dataset* Gabungan. Percobaan skenario 7 ini tidak menggunakan proses *Resampling*.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 66.25% dengan menggunakan arsitektur LSTM dan CNN 1D. Hasil akurasi tertinggi kedua adalah 63.75% dengan menggunakan arsitektur GRU.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan oleh *traits Openness* dengan menggunakan arsitektur CNN 1D dan *Extraversion* dengan menggunakan arsitektur dengan menggunakan arsitektur LSTM.

* + - * 1. **Skenario 8 (Percobaan dengan menggunakan *Undersampling*)**

**Tabel 4.35** Hasil Percobaan Skenario 8 dengan Proses *Undersampling*

| Arsitektur | O | C | E | A | N |
| --- | --- | --- | --- | --- | --- |
| MLP | 72.00% | 54.55% | 64.15% | **62.16%** | 67.80% |
| LSTM | 52.00% | 40.26% | 45.28% | 48.65% | 49.15% |
| GRU | 52.00% | 50.65% | 60.38% | 59.46% | 45.76% |
| CNN 1D | **78.00%** | **58.44%** | **79.25%** | 52.70% | **74.58%** |
| CNN 1D+LSTM | 72.00% | 51.95% | 77.36% | 51.35% | 71.19% |

Tabel 4.35 di atas merupakan percobaan dengan menggunakan *dataset* Gabungan. Percobaan skenario 8 ini menggunakan proses *Undersampling*.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 79.25% dengan menggunakan arsitektur CNN 1D. Hasil akurasi tertinggi kedua adalah 78.00% dengan menggunakan arsitektur CNN 1D. Hasil skenario ini menunjukkan penggunaan *Undersampling* pada *dataset* Gabungan ini berhasil meningkatkan hasil akurasi secara signifikan.

Untuk hasil masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan pada *traits Extraversion* dengan menggunakan arsitektur CNN 1D. Hasil yang didapatkan pada skenario ini menunjukkan terdapat 4 *traits* yang mengalami peningkatan akurasi yaitu *traits Openness* dengan 78.00% meningkat 11.75% pada skenario sebelumnya, *traits Extraversion* dengan 79.25% meningkat 13.00% pada skenario sebelumnya, *traits Agreeableness* dengan 62.16% meningkat 4.66% pada skenario sebelumnya, dan *traits Neuroticism* dengan 74.58% meningkat 10,83% pada skenario sebelumnya. Hanya terdapat 1 *traits* yang mengalami penurunan akurasi yaitu *traits Conscientiousness* dengan 58.44% menurun 0.31% pada skenario sebelumnya.

* + - * 1. **Skenario 9 (Percobaan dengan menggunakan *Oversampling*)**

**Tabel 4.36** Hasil Percobaan Skenario 9 dengan Proses *Oversampling*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Arsitektur | O | C | E | A | N |
| MLP | 46.15% | 45.56% | **63.21%** | 45.12% | 52.04% |
| LSTM | 48.08% | 44.57% | 51.92% | 48.75% | 50.96% |
| GRU | 52.88% | **57.61%** | 50.96% | 51.25% | 49.04% |
| CNN 1D | 49.04% | 48.89% | 61.32% | **53.66%** | **58.16%** |
| CNN 1D+LSTM | **54.81%** | 54.35% | 48.08% | 47.50% | 56.73% |

Tabel 4.36 di atas merupakan percobaan yang menggunakan *dataset* Gabungan. Percobaan skenario 9 ini menggunakan proses *Oversampling*.

Hasil akurasi terbaik yang berhasil didapatkan pada skenario ini adalah 63.21% dengan menggunakan arsitektur MLP. Hasil akurasi tertinggi kedua adalah 58.16% dengan menggunakan arsitektur CNN 1D. Hasil skenario ini menunjukkan penggunaan *Oversampling* pada *dataset* Gabungan ini tidak berhasil meningkatkan hasil akurasi.

Untuk masing-masing *traits* kepribadian, hasil akurasi tertinggi didapatkan pada *traits Extraversion* dengan menggunakan arsitektur MLP. Hasil yang didapatkan pada skenario ini menunjukkan bahwa tidak ada *traits* yang mengalami peningkatan akurasi dari skenario sebelumnya.

* + - * 1. **Kesimpulan Hasil Implementasi *Deep learning* pada *Dataset* Gabungan**

**Tabel 4.37** Kesimpulan hasil implementasi *Deep learning* pada *Dataset* Gabungan

|  |  |  |  |
| --- | --- | --- | --- |
| ***Traits*** | **Accuracy** | ***Algorithm*** | **Skenario** |
| ***Openness*** | 78.00% | CNN 1D | Skenario 8 |
| ***Conscientiousness*** | 58.75% | CNN 1D | Skenario 7 |
| ***Extraversion*** | 79.25% | CNN 1D | Skenario 8 |
| ***Agreeableness*** | 62.16% | MLP | Skenario 8 |
| ***Neuroticism*** | 74.58% | CNN 1D | Skenario 8 |

Implementasi *Deep learning* pada *dataset* gabungan ini berhasil mendapatkan akurasi tertinggi dengan menerapkan proses *Undersampling* seperti yang terjadi pada percobaan dengan *dataset* sebelumnya. *Traits* *Openness*, *Extraversion*, *Agreeableness*, dan *Neuroticism* mendapatkan akurasi tertinggi dengan menggunakan proses *Undersampling*. Arsitektur yang paling dominan adalah CNN 1D yang mendapatkan akurasi tertinggi di semua *traits* kecuali *Agreeableness*. Akurasi tertinggi yang berhasil didapatkan adalah 79.25% untuk *traits* *Extraversion* dengan arsitektur CNN 1D dan penggunaan proses *Resampling*. Hasil kesimpulan dari implementasi *Deep learning* pada *dataset* gabungan ini dapat dilihat pada Tabel 4.37.

* + - 1. **Kesimpulan Implementasi *Deep learning***

**Tabel 4.38** Kesimpulan Hasil Implementasi *Deep learning*

|  |  |  |  |
| --- | --- | --- | --- |
| ***Traits*** | **Accuracy** | **Arsitektur** | **Proses** |
| ***Openness*** | 78.00% | CNN 1D | *Undersampling* |
| ***Conscientiousness*** | 68.00% | CNN 1D | *Undersampling* |
| ***Extraversion*** | 93.33% | CNN 1D + LSTM | *Undersampling* |
| MLP |
| ***Agreeableness*** | 70.37% | MLP | *Undersampling* |
| ***Neuroticism*** | 80.00% | CNN 1D + LSTM | *Undersampling* |

Setelah melakukan implementasi metode *Deep learning* pada tiga *dataset* berbeda dengan 3 skenario di masing-masing *dataset* tersebut. Peneliti membuat sebuah kesimpulan yang diperoleh dari seluruh percobaan tersebut. Hasil akurasi tertinggi dari setiap skenario dibandingkan dan dipilih yang memiliki nilai tertinggi. Tabel hasil kesimpulan akurasi tertinggi di implementasi *Deep learning* dapat dilihat pada Tabel 4.38 di atas.

Dari hasil diatas, penggunaan proses *Undersampling* terbukti dapat meningkatkan hasil akurasi di metode *Deep learning*. Penerapan proses *Undersampling* berhasil mencatatkan akurasi tertinggi di seluruh *traits* dalam percobaan di seluruh skenario dengan *dataset* yang berbeda. *Traits* *Extraversion* memiliki akurasi tertinggi yaitu 93.33% dengan menggunakan MLP atau CNN 1D+LSTM. *Traits* *Neuroticism* memiliki akurasi tertinggi kedua dengan 80.00% menggunakan CNN 1D+LSTM. Kemudian *traits* *Openness* dengan menggunakan CNN 1D berhasil mendapatkan akurasi tertinggi 78.00%. *Traits* *Agreeableness* dengan akurasi tertinggi 70.37% menggunakan arsitektur MLP. *Traits* *Neuroticism* memiliki akurasi terendah dalam tabel kesimpulan ini dengan nilai 68.00% menggunakan CNN 1D.

**Gambar 4.2** Evaluasi perbandingan hasil implementasi *deep learning*

untuk setiap *traits* kepribadian

Dari Gambar 4.2 di atas dapat dilihat bahwa *traits* *extraversion* masih menjadi traits yang memiliki akurasi tertinggi dalam sistem penelitian ini dibandingkan dengan *traits* lainnya. Beberapa faktor yang mungkin menjadi penyebabnya adalah banyaknya status user yang mengandung kata-kata yang berkorelasi positif dengan *traits extraversion.* Hal lainnya yang masih menjadi asumsi dari peneliti adalah bahwa *traits extraversion* merupakan *traits* yang paling banyak ditunjukkan secara tidak langsung oleh para user di jejaring sosial. Untuk arsitektur yang digunakan dalam penelitian ini rata-rata memiliki akurasi yang cukup seimbang, namun arsitektur CNN 1D cukup dominan jika dibandingkan dengan yang lainnya.

Setelah hasil kedua implementasi yaitu implementasi *machine learning* dan implementasi *deep learning* didapatkan, perbandingannya dapat dilihat pada Gambar 4.3. Secara grafik dapat dilihat bahwa ternyata di semua *traits* kepribadian, arsitektur *deep learning* dapat mengungguli implementasi algoritma *machine learning* tradisional. Berbagai faktor yang membuat deep learning dapat mengungguli algoritma *machine learning* tradisional yang masih berupa asumsi dari peneliti yaitu arsitektur *deep learning* yang menggunakan *open vocabulary* memang lebih baik dalam proses *learning* dan klasifikasi linguistik serta memiliki potensi untuk menggantikan penggunaan machine learning tradisional yang selama ini digunakan dalam penelitian. Kedua, dataset yang dikumpulkan memiliki status yang tidak sesuai dengan fitur-fitur yang digunakan untuk implementasi algoritma *machine learning* tradisional.

Berdasarkan perbandingan kedua implementasi tersebut, maka peneliti mengambil keputusan untuk menggunakan implementasi *deep learning* di dalam sistem prediksi kepribadian pada penelitian ini karena memiliki akurasi yang lebih tinggi untuk semua *traits* dibandingkan dengan implementasi *machine learning*. Namun, sistem prediksi penelitian ini tidak akan menggunakan arsitektur yang sama untuk semua *traits* kepribadian, melainkan menggunakan arsitektur yang berbeda dan memiliki performa terbaik untuk masing-masing traits seperti yang dapat dilihat pada Tabel 4.38.

**Gambar 4.3** Perbandingan teknik implementasi algoritma tradisional *machine learning* dan *deep learning* berdasarkan akurasi tertinggi setiap *traits*

### Tampilan Layar Aplikasi

Hasil akhir aplikasi sistem prediksi kepribadian ini akan berupa sebuah aplikasi web. Gambar 4.4 hingga 4.13 adalah tampilan layar aplikasi.



**Gambar 4.4** Tampilan halaman utama aplikasi

Tampilan layar utama aplikasi memuat sebuah tombol “*Log In with Facebook*” di bagian tengah layar yang akan meminta *user* untuk *login* ke dalam akun Facebook mereka. Di bagian pojok kiri bawah layar terdapat tombol “*Privacy* *Policy*” yang akan memindahkan *user* ke halaman *Privacy* *Policy* yang dapat dilihat pada Gambar 4.8 dan 4.9.

Setelah *user* masuk ke dalam akun Facebook mereka, halaman akan berpindah ke bagian *Result* yang dapat dilihat pada Gambar 4.5 hingga 4.11. Halaman *Result* berisi hasil kepribadian dari *user* tersebut. Seperti terlihat pada Gambar 4.5, di mana terdapat 5 *traits* kepribadian *Big Five Personality* di bagian kanan halaman. Masing-masing *traits* tersebut akan memiliki *bar chart* yang terhubung di sisi kirinya. Setiap *bar chart* dalam keadaan tertutup. Untuk membuka *bar chart* tersebut, *user* harus melakukan *hover* ke *bar chart* tersebut atau *hover* ke masing-masing *traits.* Gambar 4.5 di bawah ini memperlihatkan keadaan di mana *user* sedang melakukan *hover* ke bagian *Openness*.



**Gambar 4.5** Tampilan halaman hasil dengan Openness bar chart terbuka



**Gambar 4.6** Tampilan halaman hasil dengan Conscientiousness bar chart terbuka

Gambar 4.6 diatas memperlihatkan keadaan ketika *user* melakukan hover di *traits* *Conscientiousness*.

****

**Gambar 4.7** Tampilan halaman hasil dengan Extraversion bar chart terbuka

Gambar 4.7 diatas memperlihatkan keadaan ketika *user* melakukan *hover* di *traits* *Extraversion*.

****

**Gambar 4.8** Tampilan halaman hasil dengan Agreeableness bar chart terbuka

Gambar 4.8 diatas memperlihatkan keadaan ketika *user* melakukan *hover* di *traits* *Agreeableness*.

****

**Gambar 4.9** Tampilan halaman hasil dengan Neuroticism bar chart terbuka

Gambar 4.9 diatas memperlihatkan keadaan ketika *user* melakukan *hover* di *traits* *Neuroticism*.



**Gambar 4.10** Tampilan semua hasil traits secara keseluruhan

****

**Gambar 4.11** Tampilan halaman hasil bagian penjelasan hasil masing-masing traits

Gambar 4.10 memperlihatkan semua hasil *traits* secara keseluruhan beserta tingkat kepribadiannya, jika hasil *traits* di bawah 50% maka *traits* dikategorikan “Low”, jika hasil *traits* di atas 50% maka *traits* dikategorikan “High”. Gambar 4.11 memperlihatkan hasil halaman *Result* setelah di *scroll* ke bawah dan berada tepat di bawah *bar chart* masing-masing *traits* sebelumnya. Tampilan di atas adalah penjelasan lebih jelas mengenai masing-masing *traits* yang diperoleh oleh *user*. Setiap *traits* akan berada dalam sebuah kotak yang berisi judul berupa *traits* contohnya *Openness* dan diikuti persentase kepribadiannya. Di bawah judul terdapat penjelasan lengkap mengenai *traits* *user* tersebut.

****

**Gambar 4.12** Tampilan halaman privacy policy

Gambar 4.12 adalah tampilan halaman *privacy* *policy* yang dapat dibuka dengan mengklik tombol “*Privacy* *Policy*” yang ada pada halaman utama. Halaman ini berisi penjelasan mengenai data apa saja yang didapatkan dari *user*, cara mendapatkan data dari *user*, apa tujuan peneliti menggunakan data tersebut, bagaimana peneliti menjaga keamanan data tersebut dan sebagainya. Di bagian paling bawah dari halaman ini juga terdapat informasi kontak dari peneliti yang dapat di lihat pada Gambar 4.13.

****

**Gambar 4.13** Tampilan halaman lanjutan privacy policy

## Evaluasi

### Evaluasi Subjektif

Evaluasi Subjektif dilakukan dengan melakukan wawancara terhadap beberapa responden yang diminta untuk mencoba dan mengevaluasi sistem prediksi kepribadian dari penelitian ini. Responden yang mengikuti wawancara merupakan pengguna aktif sosial media Facebook dan memiliki status di *feed* akun mereka.

Responden akan diberikan beberapa pertanyaan secara langsung melalui tanya jawab oleh peneliti, kemudian responden akan diminta untuk memberikan nilai kepuasan terhadap penggunaan sistem prediksi kepribadian yang telah dibuat serta saran atau masukkan oleh para Responden terhadap aplikasi. Berikut adalah informasi demografis para Responden:



**Gambar 4.14** *Pie chart* Distribusi Jenis Kelamin Responden



**Gambar 4.15** *Pie chart* Distribusi Umur Responden

Gambar 4.11 menampilkan distribusi jenis kelamin dari responden penelitian. Sedangkan Gambar 4.12 menampilkan distribusi umur dari responden. Berikut adalah poin-poin penting yang didapatkan melalui hasil wawancara terhadap para Responden.

1. 8 dari 10 responden menggunakan sosial media Facebook setiap hari. Sedangkan 2 lainnya menggunakan Facebook beberapa hari sekali.
2. Para responden berpendapat bahwa tampilan layar aplikasi sudah cukup bagus dengan nilai rata-rata 8.6 dari 10.
3. Para responden berpendapat sistem prediksi yang dikembangkan menghasilkan akurasi dalam waktu yang cepat.
4. Para responden berpendapat hasil prediksi untuk *traits* *Opennes* lumayan akurat dengan nilai rata-rata 7.6 dari 10.
5. Para responden berpendapat hasil prediksi untuk *traits* *Conscientiousness* cukup akurat dengan nilai rata-rata 7.2 dari 10.
6. Para responden berpendapat hasil prediksi untuk *traits* *Extraversion* cukup akurat dengan nilai rata-rata 6.6 dari 10.
7. Para responden berpendapat hasil prediksi untuk *traits* *Agreeableness* lumayan akurat dengan nilai rata-rata 7.5 dari 10.
8. Para responden berpendapat hasil prediksi untuk *traits* *Neuroticism* cukup akurat dengan nilai rata-rata 6.9 dari 10.
9. 8 dari 10 responden berpendapat bahwa sistem prediksi kepribadian akan bermanfaat dalam kehidupan mereka.
10. Para responden berpendapat bahwa keseluruhan sistem aplikasi sistem prediksi kepribadian ini sudah cukup baik, beberapa responden memberikan saran untuk meningkatkan hasil akurasi prediksi kepribadian, karena masih ada beberapa prediksi yang dianggap kurang tepat. Selain itu, terdapat juga saran dari responden untuk membuat versi *mobile* agar lebih mudah digunakan.

**Tabel 4.39** Penilaian Kuantitatif untuk setiap *Traits* oleh Responden Penelitian

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Responden** | ***Openness*** | ***Conscientiousness*** | ***Extraversion*** | ***Agreeableness*** | ***Neuroticism*** |
| **1** | 10 | 7 | 1 | 8 | 5 |
| **2** | 8 | 8 | 1 | 8 | 8 |
| **3** | 6 | 8 | 7 | 6 | 7 |
| **4** | 7 | 9 | 9 | 9 | 5 |
| **5** | 9 | 8 | 6 | 10 | 7 |
| **6** | 8 | 6 | 5 | 9 | 7 |
| **7** | 4 | 8 | 10 | 3 | 8 |
| **8** | 9 | 3 | 8 | 9 | 8 |
| **9** | 7 | 7 | 9 | 10 | 6 |
| **10** | 8 | 8 | 10 | 3 | 8 |
| **Rata-rata** | **7.6** | **7.2** | **6.6** | **7.5** | **6.9** |

### Evaluasi Objektif

Evaluasi Objektif dilakukan dengan membandingkan hasil yang diperoleh oleh penelitian ini dengan penelitian sebelumnya yang menggunakan model kepribadian *Big Five* *Traits* dan mencantumkan hasil akurasi penelitian, karena aspek utama yang menjadi perbandingan adalah hasil akurasi dari sistem prediksi kepribadian untuk setiap *traits* dari *Big Five Model Personality*.

* + - 1. **Sistem Prediksi Kepribadian “The *Big Five* *Traits*” Dari Data Twitter**

Penelitian yang bertujuan untuk membangun sebuah sistem prediksi kepribadian ini dilakukan oleh Wijaya, Prasetia, Febrianto, dan Suhartono. Walaupun penelitian menggunakan *dataset* dari sosial media Twitter, evaluasi dilakukan karena metode, fitur dan algoritma yang digunakan memiliki beberapa kesamaan dan pantas untuk dilakukan perbandingan hasil akurasinya. Selain itu, penelitian yang dilakukan oleh Wijaya, Prasetia, Febrianto, dan Suhartono ini juga menggunakan bahasa Inggris. Algoritma yang digunakan pada penelitian adalah Support Vector Machine, Naïve Bayes, dan ZeroR. Fitur Linguistik yang digunakan yaitu gabungan dari fitur LIWC dan MRC.

**Tabel 4.40** Perbedaan antara penelitian Sistem Prediksi Kepribadian “The Big Five *Traits*” Dari Data Twitter dan Penelitian Ini

|  |  |  |
| --- | --- | --- |
| **Perbedaan** | **Sistem Prediksi Kepribadian “The Big Five *Traits*” Dari Data Twitter** | **Penelitian Ini** |
| Five Factor Model |  |  |
| *Machine learning* |  |  |
| Support Vector Machine |  |  |
| ZeroR |  |  |
| Naïve Bayes |  |  |
| Linear Discriminant Analysis |  |  |
| Gradient Boosting |  |  |
| Logistic Regression |  |  |
| *Deep learning* |  |  |
| *Features* |  |  |
| LIWC |  |  |
| MRC |  |  |
| SPLICE |  |  |
| Open Vocabulary |  |  |

Evaluasi objektif dilakukan dengan membandingkan akurasi tertinggi dari setiap *traits* kepribadian antara kedua penelitian diatas. Penelitian yang dilakukan oleh Wijaya, Prasetia, Febrianto, dan Suhartono hanya menampilkan hasil akurasi dengan algoritma SVM karena algoritma tersebut merupakan fokus dari penelitian dan digunakan pada sistem prediksi. Perbandingan akurasi kedua penelitian dapat dilihat pada Tabel 4.41.

Hasil perbandingan antara metode *Machine learning* kedua penelitian cukup seimbang. Penelitian ini hanya berhasil mengungguli hasil akurasi penelitian sebelumnya pada *traits* *Openness*, *Extraversion*. Sedangkan untuk *traits* *Conscientiousness*, *Agreeableness*, dan *Neuroticism*, masih belum dapat mengungguli akurasi dari penelitian sebelumnya. Namun, penggunaan implementasi *Deep learning* penelitian ini memiliki hasil yang lebih baik dan unggul di 3 *traits* kepribadian yaitu, *Openness*, *Extraversion*, dan *Agreeableness*.

Hasil akurasi yang berbeda kemungkinan besar disebabkan karena *dataset* yang digunakan berbeda dalam jumlah maupun jenis. *Dataset* yang digunakan pada penelitian sebelumnya sebanyak kurang lebih 5 ribu data, sedangkan penelitian ini hanya menggunakan sekitar 400 *dataset*. Penggunaan metode *Deep learning* memberikan hasil yang cukup baik meski dengan *dataset* yang lebih sedikit dan mungkin dapat ditingkatkan lagi jika diimplementasikan dengan *dataset* yang lebih besar.

**Tabel 4.41** Tabel perbandingan akurasi penelitian Sistem Prediksi Kepribadian “The Big Five *Traits*” Dari Data Twitter dan Penelitian Ini

|  |  |  |  |
| --- | --- | --- | --- |
| ***Traits*** | **Sistem Prediksi Kepribadian “The Big Five *Traits*” Dari Data Twitter** | **Penelitian Ini** | |
| ***Machine Learning*** | ***Machine Learning*** | ***Deep Learning*** |
| ***Openness*** | 61.436 (SVM) | 70.4  (Logistic Regression) | 78 (CNN 1D) |
| ***Conscientiousness*** | 80.876 (SVM) | 67.33  (LDA) | 68 (CNN 1D) |
| ***Extraversion*** | 64.3938 (SVM) | 79.33  (LDA) | 93.33 (CNN 1D + LSTM) |
| ***Agreeableness*** | 77.8957 (SVM) | 63.2  (Gradient Boosting) | 70.37 (MLP) |
| ***Neuroticism*** | 70.5351 (SVM) | 70  (Naïve Bayes) | 80 (CNN 1D + LSTM) |

* + - 1. ***Personality Traits Recognition on Social Network – Facebook***

Penelitian sebelumnya yang menggunakan sosial media Facebook sebagai *dataset* dilakukan oleh Alam, Stepanov, dan Riccardi. Penelitian ini menggunakan *dataset* yang sama persis dengan penelitian kali ini yaitu *dataset* my*Personality* (Kosinski, 2015) sebanyak 250 data *user* dan kurang lebih 10.000 status. Algoritma yang digunakan pada penelitian oleh Alam, Stepanov, dan Riccardi adalah SVM, Logistic Regression, dan Naïve Bayes. Feature Extraction dan klasifikasi pada penelitian sebelumnya menggunakan WEKA (Witten, 2011).

**Tabel 4.42** Perbedaan antara penelitian *Personality Traits Recognition on Social Network – Facebook* dan Penelitian Ini

|  |  |  |
| --- | --- | --- |
| **Perbedaan** | ***Personality Traits Recognition on Social Network – Facebook*** | **Penelitian Ini** |
| Five Factor Model |  |  |
| *Machine learning* |  |  |
| Support Vector Machine |  |  |
| Naïve Bayes |  |  |
| Logistic Regression |  |  |
| Linear Discriminant Analysis |  |  |
| Gradient Boosting |  |  |
| *Deep learning* |  |  |
| *Features* |  |  |
| LIWC |  |  |
| SPLICE |  |  |
| Open Vocabulary |  |  |

Evaluasi objektif dilakukan dengan membandingkan akurasi tertinggi dari setiap *traits* kepribadian antara penelitian yang dilakukan oleh Alam, Stepanov, dan Riccardi dengan penelitian ini. Perbedaan antara kedua penelitian terdapat pada fitur yang dipakai pada implementasi *Machine learning* di mana penelitian sebelumnya mengimplementasikan open-vocabulary sebagai fitur, sedangkan penelitian ini menggunakan closed-vocabulary. Walaupun demikian, implementasi *Deep learning* yang digunakan pada penelitian ini juga menggunakan open-vocabulary sebagai fitur utama. Untuk itu hasil akurasi dengan menggunakan implementasi *Deep learning* juga akan diikutsertakan dalam tabel perbandingan hasil akurasi. Hasil perbandingan akurasi kedua penelitian dapat dilihat pada Tabel 4.43 di bawah ini.

**Tabel 4.43** Tabel perbandingan akurasi penelitian *Personality Traits Recognition on Social Network – Facebook* dan Penelitian Ini

|  |  |  |  |
| --- | --- | --- | --- |
| ***Traits*** | ***Personality Traits Recognition on Social Network – Facebook*** | **Penelitian Ini** | |
| ***Machine Learning*** | ***Machine Learning*** | ***Deep Learning*** |
| ***Openness*** | 69.48 (Naïve Bayes) | 70.4  (Logistic Regression) | 78 (CNN 1D) |
| ***Conscientiousness*** | 59.4 (Naïve Bayes) | 67.33  (LDA) | 68 (CNN 1D) |
| ***Extraversion*** | 58.6 (Naïve Bayes) | 79.33  (LDA) | 93.33 (CNN 1D + LSTM) |
| ***Agreeableness*** | 59.16 (Naïve Bayes) | 63.2  (Gradient Boosting) | 70.37 (MLP) |
| ***Neuroticism*** | 63.0 (Naïve Bayes) | 70  (Naïve Bayes) | 80 (CNN 1D + LSTM) |

Dari Tabel 4.43 dapat dilihat bahwa implementasi *Machine learning* yang digunakan pada penelitian ini mengungguli hasil akurasi dari semua *traits* pada penelitian sebelumnya. Fitur yang digunakan merupakan salah satu alasan penting yang menyebabkan perbedaan hasil kedua penelitian. Namun, terdapat 3 *traits* di mana penelitian ini mengungguli penelitian sebelumnya, menggunakan algoritma yang tidak digunakan pada penelitian sebelumnya yaitu Linear Discriminant Analysis (LDA) dan Gradient Boosting. Hal itu dapat menjadi faktor keunggulan penelitian ini.

Jika dibandingkan dengan penggunaan fitur yang sama yaitu open-vocabulary. Penelitian ini yang menggunakan *Deep learning* jauh mengungguli hasil penelitian sebelumnya di semua *traits* kepribadian. Hal ini menghasilkan kesimpulan bahwa open-vocabulary dapat digunakan lebih efektif pada metode *Deep learning*. Semakin luas dan besar fitur dari *dataset* yang dimiliki, semakin efektif metode *Deep learning* bekerja.

* + - 1. ***Deep Learning-Based Document Modeling for Personality Detection from Text***

Salah satu penelitian yang menerapkan implementasi *Deep Learning* dilakukan oleh Majumder, Poria, Gelbukh, & Cambria. Jenis kepribadian yang digunakan adalah *Big Five Personality Traits* yang sama dengan penelitian ini. Metode yang digunakan oleh penelitian Majumder et al. meliputi *preprocessing* data inputdan *filtering*, *Feature Extraction*, dan *Classification*. Untuk *Word-Level Feature Extraction*, penelitian oleh Majumder et al. menggunakan *word2vec embeddings* dan klasifikasi menggunakan arsitektur CNN dan classifier MLP serta SVM. *Dataset* yang digunakan dalam penelitian sebelumnya terdiri dari 2.468 esai tanpa nama yang telah diberi label kepribadian dari penulis, Untuk melakukan evaluasi terhadap hasil *training*. Penelitian sebelumnya menggunakan *10-fold cross-validation*. Perbandingan jelas antara perbedaan metodologi antara penelitian sebelumnya dengan penelitian ini dapat dilihat pada Tabel 4.44.

**Tabel 4.44** Perbedaan antara penelitian *Deep Learning-Based Document Modeling for Personality Prediction from Text* dan Penelitian Ini

| **Perbedaan** | *Deep Learning-Based Document Modeling for Personality Prediction from Text* | **Penelitian Ini** |
| --- | --- | --- |
| Five Factor Model |  |  |
| *Machine learning* |  |  |
| SVM |  |  |
| Naïve Bayes |  |  |
| Logistic Regression |  |  |
| *Deep learning* |  |  |
| CNN |  |  |
| Mairesse |  |  |
| MLP |  |  |
| LSTM |  |  |
| GRU |  |  |
| *Features* |  |  |
| LIWC |  |  |
| SPLICE |  |  |
| Open Vocabulary |  |  |

Evaluasi objektif dilakukan dengan membandingkan akurasi tertinggi dari setiap *traits* kepribadian antara penelitian yang dilakukan oleh Majumder, Poria, Gelbukh, & Cambria dengan penelitian ini. Kedua penelitian sama-sama menggunakan implementasi *Deep Learning* dan arsitektur yang sama yaitu CNN. Perbedaan terdapat pada jenis *dataset* yang digunakan di mana penelitian sebelumnya menggunakan *dataset* esai sebanyak 2,468 esai, sedangkan penelitian ini menggunakan *dataset* status pengguna sosial media Facebook yang berjumlah sekitar 400 user. Hasil perbandingan akurasi kedua penelitian dapat dilihat pada Tabel 4.45 di bawah ini.

**Tabel 4.45** Tabel perbandingan akurasi penelitian *Deep Learning-Based Document Modeling for Personality Prediction from Text* dan Penelitian Ini

| ***Traits*** | ***Deep Learning-Based Document Modeling for Personality Detetction from Text*** | **Penelitian Ini** |
| --- | --- | --- |
| ***Deep Learning*** | ***Deep Learning*** |
| ***Openness*** | 62.68 (CNN + Mairesse) | 78 (CNN 1D) |
| ***Conscientiousness*** | 57.30 (CNN + Mairesse) | 68 (CNN 1D) |
| ***Extraversion*** | 58.09 (CNN + Mairesse) | 93.33 (CNN 1D + LSTM) |
| ***Agreeableness*** | 56.71 (CNN + Mairesse) | 70.37 (MLP) |
| ***Neuroticism*** | 59.38 (CNN + Mairesse) | 80 (CNN 1D + LSTM) |

Dari tabel perbandingan akurasi di atas dapat dilihat bahwa penelitian ini yang mengimplementasikan *Deep Learning* untuk sistem prediksi kepribadian berhasil mengungguli penelitian sebelumnya di semua *traits* kepribadian. Penelitian sebelumnya memiliki akurasi tertinggi yaitu 62.68% untuk *traits* *Openness*, sedangkan penelitian ini berhasil unggul sekitar 16% untuk *traits* yang sama. Arsitektur kedua penelitian ini juga di dominasi oleh arsitektur CNN dan gabungan dengan LSTM. Perbedaan akurasi kemungkinan besar disebabkan oleh *dataset* yang berbeda.