

Daftar Tabel

Table 1. Social Feature

| Social Feature | Descriptions |
|---------------------------|---|
| Number of Followers | The number of followers that the user has |
| Number of Following | The number of users following |
| Number of Tweets | The number of users' tweets |
| Number of URL | The large number of URLs shared by users |
| Number of Media URL | The large number of media URLs shared by users |
| Number of Hashtags | The number of users' hashtags |
| Number of Retweets | The number of user retweets |
| Number of Mentions | The number of user mentions |
| Number of Capital Letters | The number of uppercase letters used by Twitter users |

Table 2. Hasil Preprocessing

| Id | Label | Akun | Tweet |
|-----|---------------|--------------|---|
| 1 | Openness | 12290F | its besar salah for me butuh kj to mtk... |
| 2 | Openness | Abiyoga | if you like pineapple and you like pizza you should... |
| 3 | Openness | Adityarestya | mbak analogi nya ting logika pikir nya lebih teliti mana wkwk... |
| ... | ... | ... | ... |
| 315 | Agreeableness | teajusbiasa_ | astagfirullah mudafokus ning mburiii jebul sing harap koyo aku... |

Table 3. Contoh LIWC

| Username | Tweet | Orang pertama |
|--------------|---|---------------|
| ShiroiRayn | ['aku', 'dek'] | -0,19 |
| _mericabubuk | ['kali', 'dilaranglarang', 'mas', 'aku', 'tunggu', 'tungguuuu'] | -0,19 |

Table 4. Akurasi Rasio

| Ratio | Accuracy (%) |
|--------------|--------------|
| 90:10 | 37,50 |
| 80:20 | 38,10 |
| 70:30 | 37,89 |
| 60:40 | 34,13 |
| 50:50 | 29,75 |

Table 5. Akurasi Skenario 1

| Conditions | Accuracy (%) |
|--------------------------------|--------------|
| Logistic Regression (Baseline) | 38,10 |
| Baseline + K-Means SMOTE | 42,11 |

Table 6. Akurasi Skenario 2

| Conditions | Accuracy (%) |
|------------------------------------|--------------|
| Baseline + K-Means SMOTE | 42,11 |
| Baseline + K-Means SMOTE + LIWC | 52,00 |

Table 7. Akurasi Skenario 3

| Conditions | Accuracy (%) |
|---|--------------|
| Baseline + K-Means SMOTE + LIWC | 52,00 |
| Baseline + K-Means SMOTE + LIWC + IndoBERT | 75,11 |

Table 8. Akurasi Skenario 4

| Conditions | Accuracy (%) |
|--|----------------|
| Baseline | 38,10 |
| Baseline + K-Means SMOTE | 42,11 (+10,5) |
| Baseline + K-Means SMOTE + LIWC | 52,00 (+36,4) |
| Baseline + K-Means SMOTE + LIWC + IndoBERT | 75,11 (+97,1) |
| Baseline + K-Means SMOTE + LIWC + IndoBERT + ANOVA | 86,17 (+126,6) |
| Baseline + K-Means SMOTE + LIWC + IndoBERT + ANOVA + Hyperparameter | 89,57 (+135,1) |