



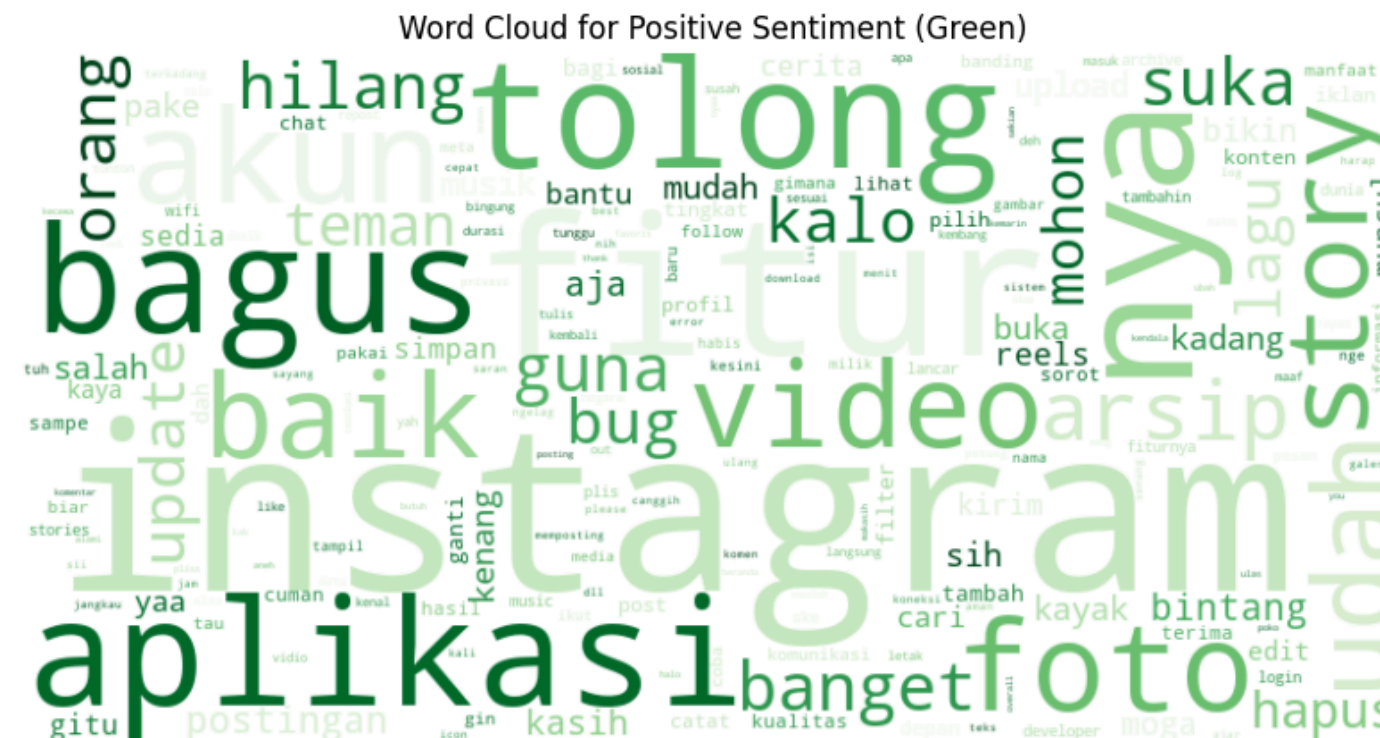
# Background

# What Happened?

[illegible]

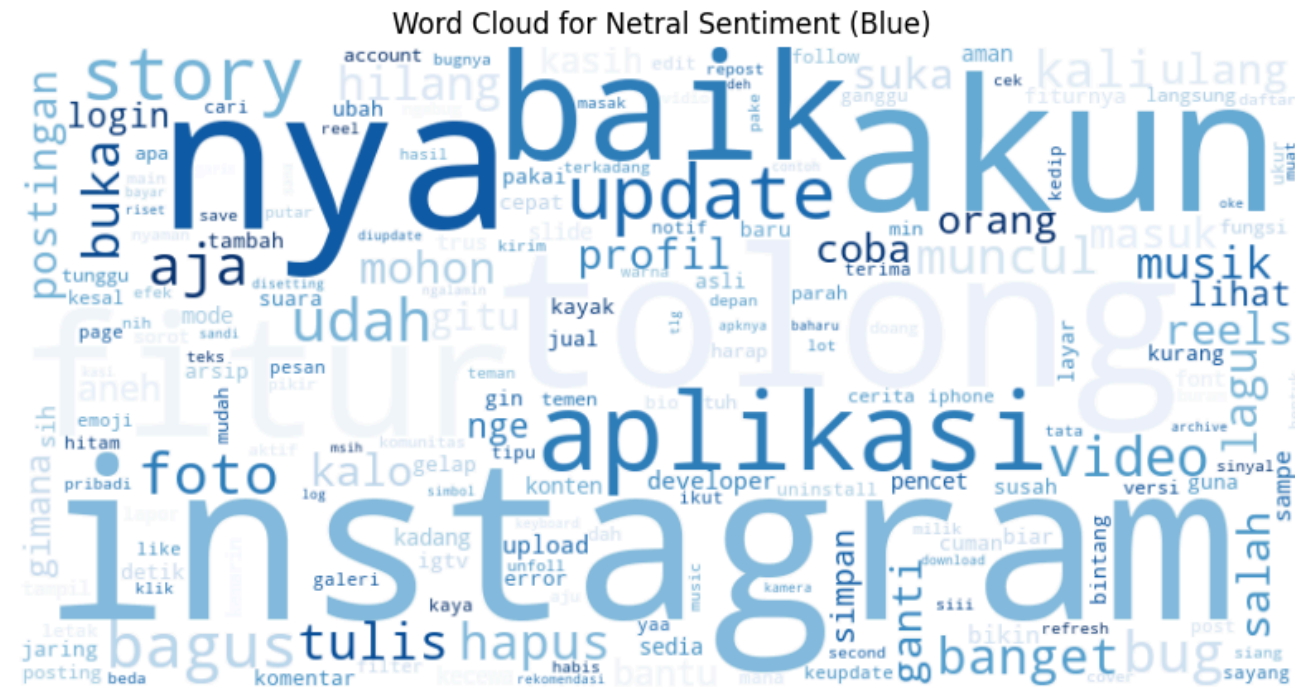
Most **users type words** such as "**saya**", "**tidak**", "**bisa**", "**dan**", "**instagram**", "**akun**", "**fitur**", and others when writing **Instagram reviews**.

# What positive words often appear in reviews?



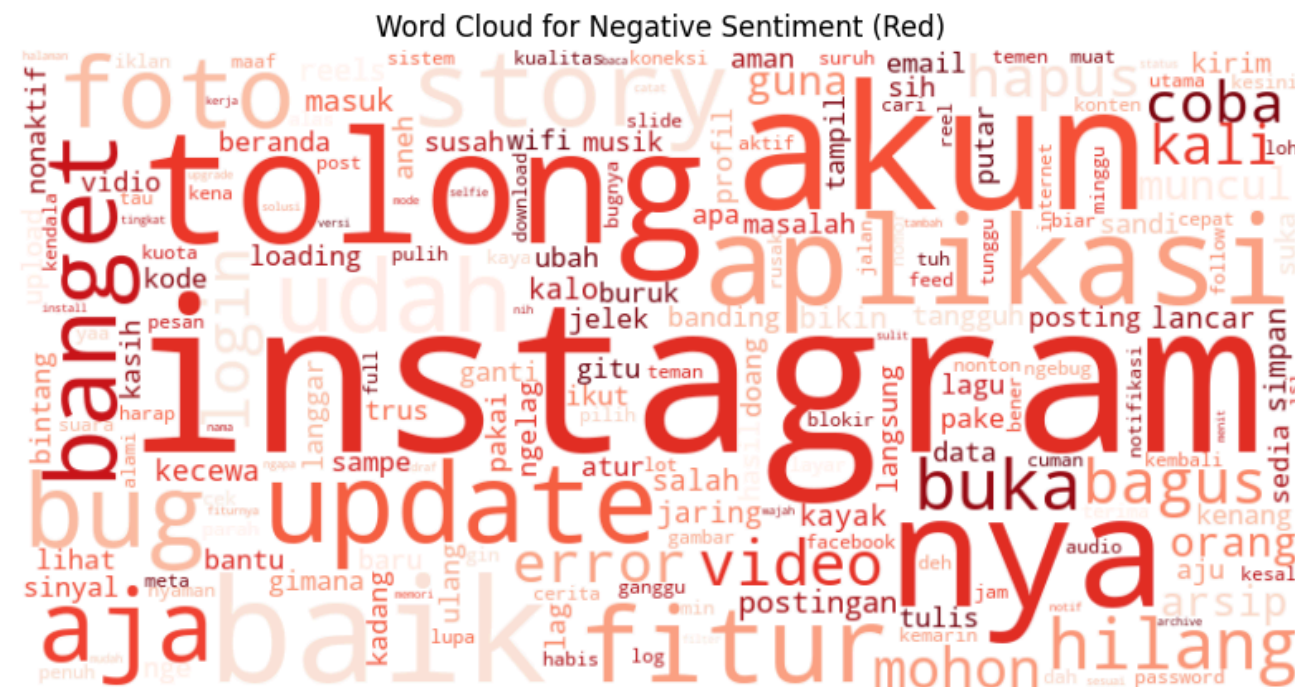
Most **users type words** such as **"bagus", "aplikasi", "instagram", "akun", "fitur",** and others when writing **Instagram reviews.**

What neutral words often appear in reviews?



Most **users type words** such as "**aplikasi**", "**baik**", "**update**", "**instagram**", "**akun**", "**fitur**", and others when writing **Instagram reviews**.

What negative words often appear in reviews?



Most **users type words** such as **"tolong", "bug", "instagram", "akun", "fitur", "buruk"** and others when writing **Instagram reviews**.

## Preprocessing

## Workflow for Preprocessing Text:



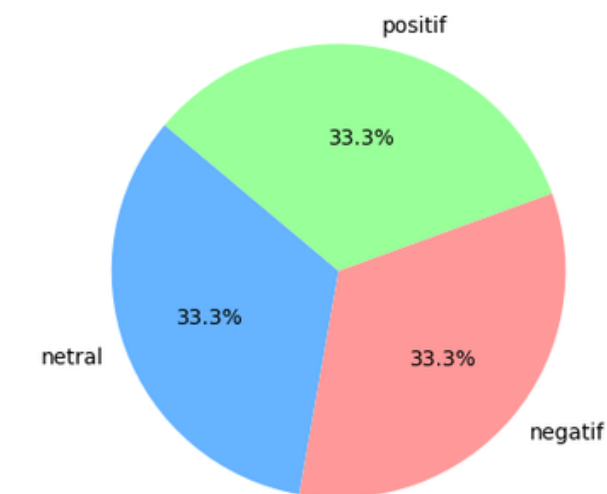
	content	Normalisasi	Tokenize	Stopword	Stemming
0	gatau ig skrng banyak bug nya mulai dari posti...	instagram sekarang banyak bug nya mulai dari p...	[instagram, sekarang, banyak, bug, nya, mulai,...	[instagram, bug, nya, postingan, suaranya, sim...	instagram bug nya postingan suara simbol speak...
1	tolong diperbaiki dong bug suara yg dipostinga...	tolong diperbaiki dong bug suara yang dipostin...	[tolong, diperbaiki, dong, bug, suara, yang, d...	[tolong, diperbaiki, bug, suara, diposting, ...	tolong baik bug suara diposting postingan in...
2	gimana ya ig makin kesini makin banyak bug nya...	gimana ya instagram makin kesini makin banyak ...	[gimana, ya, instagram, makin, kesini, makin, ...	[gimana, ya, instagram, kesini, bug, nya, makl...	gimana instagram kesi bug nya maklum bikin k...
3	kenapa ya akun instagram saya ketiga nya jika ...	kenapa ya akun instagram saya ketiga nya jika ...	[kenapa, ya, akun, instagram, saya, ketiga, ny...	[ya, akun, instagram, ketiga, nya, menonton, v...	akun instagram tiga nya tonton video reels suk...
4	saya suka intagram begitu banyak video bermanf...	saya suka intagram begitu banyak video bermanf...	[saya, suka, intagram, begitu, banyak, video, ...	[suka, intagram, video, bermanfaat, dalamnya, ...	suka intagram video manfaat dalam namu jarang ...

## Feature Extraction

I did **TF-IDF method** on feature extraction.

## Oversampling with SMOTE

### Distribution of Sentiment Categories After SMOTE



This dataset is **imbalanced**, I use **SMOTE** to make it **balanced**.



## Modeling

I use **Random Forest, Decision Tree, Logistic Regression, Naive Bayes, KNeighborsClassifier** and **Linear SVC** for Model Development.

## Evaluation

- I'm still **paying attention** to the **accuracy score** as well since this metric is easier to interpret.
- The model's performance was evaluated using metrics such as **precision, recall, and F1-score**.

	Model	Accuracy	F1_Score	Recall	Precisions
0	RandomForestClassifier	91.8%	0.918828	0.917738	0.925450
1	SVM	89.5%	0.893663	0.894602	0.894957
2	LogisticRegression	87.1%	0.870770	0.871465	0.871669
3	Naive Bayes	79.7%	0.794447	0.796915	0.812976
4	DecisionTreeClassifier	73.0%	0.729313	0.730077	0.729701
5	KNeighborsClassifier	65.6%	0.556268	0.655527	0.778645

**Random Forest Classifier give the highest performance**

## Deployment with Streamlit

To **deploy the Streamlit app**, I ensured all dependencies were installed, prepared a **requirements.txt** file, and **deployed via Streamlit Sharing**, linking a **GitHub repository** for **automated updates**.

### Sentiment Analysis Instagram Review

Input your text here:

aplikasi bagus dan keren

Analyze

Sentiment: Positif

## Conclusion

The **RandomForestClassifier** demonstrated the **best performance** with an **accuracy of 91.8%, F1-score of 0.9188, recall of 0.9177, and precision of 0.9254**. This model exhibits an excellent balance between precision and recall, making it the **top choice for sentiment classification**. It was **deployed via Streamlit**, enabling an **interactive interface for real-time sentiment predictions** (positive, negative, or neutral).