

Figure 1. Bentuk aturan pohon aplikasi.

Metode

Berdasarkan gambar(3), dengan mengubah gambar(1) menjadi gambar(2) yang merupakan bentuk CFGnya, memungkinkan sistem dapat membaca setiap aturan frasa pada data yang ada dan mengubah jenis label frasa nya ketika aturan frasa tersebut dianggap salah oleh sistem dengan label jenis frasa yang dianggap benar oleh sistem.

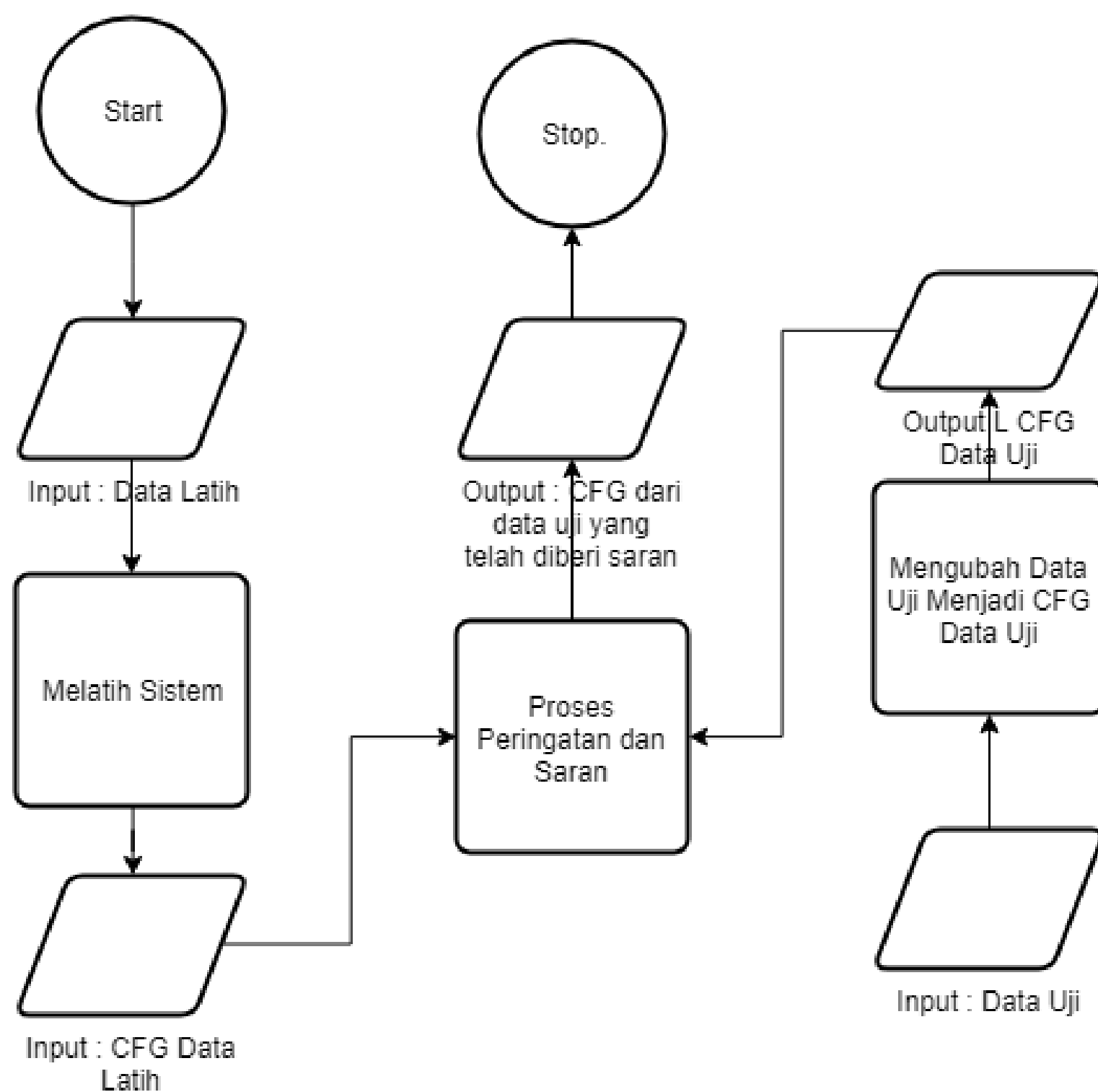


Figure 3. Flowchart sistem.

Hasil

Sistem menggunakan metode recall and precision (table(1) dan rumus(1)) untuk mendapatkan nilai precision, recall, akurasi serta F1 score dari sistem. Dari hasil yang didapat (table(2)), baik akurasi, recall serta precision dari sistem terbukti dapat bersaing dengan penilaian dari rata-rata manusia (responden) dalam menentukan label jenis frasa yang tepat pada suatu komposisi kata atau pola aturan pada kalimat.

Latar Belakang

TIKSentenceTreeMaker adalah aplikasi yang digunakan untuk membantu memberi label jenis frasa pada komposisi kata tertentu. Selain mempermudah pengguna dalam memberi label, aplikasi ini masih memiliki kekurangan seperti tidak adanya sistem peringatan dan saran ketika pengguna terdeteksi melakukan salah dalam memberi label.

NP -> NUM_2.700 NNO_penduduk
NP -> ADJ_Sebanyak NP
VP -> ADK_telah VBP_terjebak
PP -> PPO_oleh NNO_banjir
VP -> VP PP
S -> NP VP SYM_.

Figure 2. Bentuk aturan CFG.

Table 1. Tabel klasifikasi kebenaran untuk recall and precision.

Nilai Data Sebelum Dima-sukan ke Sistem	Nilai Data Setelah Dima-sukan ke Sistem	Klasifikasi
True	True	True Negative (TN)
True	False	False Negative (FN)
False	True	True Positive (TP)
False	False	False Positive (FP)

$$Recall = \frac{TP}{TP + FN} \times 100\%$$

$$Precision = \frac{TP}{TP + FP} \times 100\%$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \times 100\%$$

$$F1 \text{ Score} = 2x\left(\frac{Recall \times Precision}{Recall + Precision}\right) \times 100\%$$

Rumus 1. Rumus akurasi, recall, precision dan F1 Score.

Table 2. Tabel hasil akurasi yang didapat dengan berbagai sumber..

Data test	Accuracy	Recall	Precision	F1 Score
Lower Bound	80.00%	Na	Na	Na
Upper Bound	100.00%	100.00%	100.00%	100.00%
Data uji awal	87.88%	Na	Na	Na
Hasil saran sistem	95.84%	100.00%	65.71%	79.31%
Rata-rata responden	91.75%	67.45%	58.33%	61.97%

References

1. Samsuri. Analisis Bahasa. Erlangga, 1994.
2. Inacl pos tagging convention konvensi pelabelan kelas kata inacl/malkin. 2017.
3. G. Z. N. Totok Suhardijanto, Ayu Purwantiari. Inacl syntax parsing convention konvensi pelabelan struktur kalimat inacl/malkin. 2017.
4. Pengembangan, Badan. "Pedoman Umum Ejaan Bahasa Indonesia." (2016).
5. A. Clark, C. Fox, and S. Lappin. The Handbook of Computational Linguistics and Natural Language Pro-cessing. Blackwell Handbooks in Linguistics. Wiley, 2010.
6. A. Taylor, M. Marcus, and B. Santorini. The Penn Treebank: An Overview, pages 5–22. Springer Netherlands, Dordrecht, 2003