BAB V
KESIMPULAN

5.1 Kesimpulan

1. Berdasarkan hasil penelitian yang telah dilakukan dengan mengimplementasikan algoritma *Support Vector Machine* pada dataset mengenai kenaikan pajak pertambahan nilai 12% pada media sosial *X* dengan hasil pengambilan data sebanyak 3.332 data. Kemudian dilakukan *preprocessing* yang komprehensif dengan tahapan *Cleaning, Casefolding, Slangword, Tokenizing, Stopword Removal*, dan *Stemming*. Setelah itu dilakukan proses labeling dataset dan didapatkan sebanyak 646 atau 25% label positif dan 1936 atau 75% label negatif.
2. Pada proses klasifikasi pada penelitian ini menggunakan berbagai jenis kernel berbeda yaitu *linear, polynomial, radial basis function (RBF), dan sigmoid* serta membagi dataset kedalam tiga skenario pembagian data yaitu 70:30, 80:20, dan 90:10. Dalam skenario 70:30, *Kernel RBF* menunjukkan akurasi terbaik sebesar 0.78 atau 78%, dengan keseimbangan *precision, recall,* dan *f1-score* yang baik. Lalu pada skenario 80:20, *Kernel linear* menunjukkan performa terbaik secara keseluruhan dengan *accuracy* 0.75 atau 75%, *precision* 0.76 atau 76%, dan *f1-score* 0.73 atau 73%, skenario ini memberikan keseimbangan yang baik dalam pendekatan kedua kelas. Sedangkan dalam skenario 90:10 meskipun *kernel Polynomial* menunjukkan akurasi tinggi, *Kernel linear* tetap menunjukkan performa terbaik secara keseluruhan dengan *accuracy* 0.73 atau 73%, *precision* 0.74 atau 74%, dan *f1-score* 0.73 atau 73%.

5.2 Saran

Dari hasil penelitian yang dilakukan masih banyak kekurangan, dengan demikian peneliti berharap penelitian ini dapat dikembangkan, beberapa saran dati penulis antara lain:

1. Pada penelitian selanjutnya diharapkan dapat mengeksplorasi penggunan algoritma klasifikasi lain seperti *Naive Bayes, Long Short-*

Term Memory (LSTM), atau *Bidirectional Encoder Representations from Transformers (BERT)* untuk membandingkan kinerja dan mencari model yang lebih optimal.

2. Pada proses extraksi fitur diharapkan dapat menambahkan fitur lain seperti *FastText* untuk menangkap makna semantik dan konteks kata yang lebih kaya, yang mungkin dapat meningkatkan akurasi model.