

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan serangkaian proses penelitian yang sudah dilaksanakan, penulis menarik kesimpulan sebagai berikut:

1. Berdasarkan analisis kasus yang sudah dilakukan, model prediksi dirancang dengan beberapa praproses dan penyesuaian. Data video disample sehingga setiap data video diwakili oleh 24 citra (satu citra pada setiap jamnya). Setelah citra didapatkan, dilakukan proses labeling dengan menggunakan katalog *solar flare*. Selanjutnya citra disegmentasi dan dilakukan praproses *machine learning* dengan menggunakan PCA dan penanganan *imbalanced data*. Langkah terakhir adalah memproses data ke dalam algoritma *random forest*.
2. Dalam mengimplementasikan sistem, penggunaan bahasa pemrograman python sangat membantu proses penelitian. Alasannya, adalah karena python memiliki banyak *library* yang sangat bermanfaat selama proses pengembangan. Salah satunya adalah algoritma *pytesseract* yang digunakan penulis untuk melakukan otomatisasi pemotongan *frame* berdasarkan *timestamp* yang tersedia pada citra bergerak. Juga library lain seperti *numpy* dan *opencv* yang bermanfaat selama proses pengembangan model prediksi. Juga python memiliki *library machine learning* yang penulis gunakan sebagai pembanding luaran algoritma seperti *sklearn* dan *keras*.
3. Dari hasil eksperimen, penggunaan algoritma segmentasi *seeded region growing* mampu memberikan rerata nilai *f-measure learning* yang lebih baik jika dibandingkan dengan tidak menggunakan algoritma tersebut. Namun, memberikan waktu eksekusi algoritma *machine learning* yang jauh lebih besar pada sebagian besar algoritma. Juga penggunaan algoritma *random forest* yang dibangun memiliki luaran *f-measure* yang lebih baik jika dibandingkan dengan algoritma SVR. Namun, memiliki luaran *f-measure* yang lebih kecil jika dibandingkan dengan algoritma XGB dan MLP. Hasil algoritma *random forest* memiliki waktu eksekusi yang sangat

besar jika dibandingkan dengan algoritma *random forest* yang dibangun dengan menggunakan library *sklearn*. Perbandingan waktu eksekusi rata – rata adalah sebesar 5000x lipat. Namun, tidak memiliki perbedaan yang besar jika dilihat pada luaran *f-measure*.

5.2 Saran

Diharapkan penelitian selanjutnya dalam pembangunan sistem prediksi dengan menggunakan data SWAP PROBA2 dapat menggunakan teknik dan skema yang lebih mutakhir, baik dari algoritma segmentasi maupun *machine learning*. Adapun beberapa saran secara spesifik adalah sebagai berikut :

1. Perlu dilakukan optimasi terhadap algoritma *seeded region growing* yang dibangun karena waktu komputasi yang dibutuhkan masih sangat besar.
2. Proses OCR dilakukan terhadap citra utuh, sehingga banyak kegagalan saat melakukan deteksi *teks*. Juga, menyebabkan biaya komputasi menjadi besar. Proses OCR dapat dilakukan secara spesifik terhadap daerah yang mengandung *timestamp*. Sehingga *noise* akan lebih sedikit pada saat proses OCR. Juga biaya komputasi yang dibutuhkan menjadi lebih rendah.
3. Perlu adanya optimasi algoritma *random forest* yang dibangun secara *native* sehingga waktu eksekusi yang dibutuhkan untuk melakukan proses learning dapat diminimalisi