

**KLASIFIKASI GENRE FILM BERDASARKAN WIKIPEDIA MOVIE  
PLOTS MENGGUNAKAN NAIVE BAYES CLASSIFIER**

**TUGAS AKHIR**

Diajukan Sebagai Salah Satu Syarat untuk Menyelesaikan Program Strata-1 pada  
Jurusan Sistem Informasi Fakultas Teknologi Informasi Universitas Andalas



**JURUSAN SISTEM INFORMASI  
FAKULTAS TEKNOLOGI INFORMASI  
UNIVERSITAS ANDALAS**

**2021**

## ABSTRAK

Film telah berada lebih dari satu abad dalam kehidupan kita sehari-hari ini baik dalam bentuk digital maupun fisik. Film tersebut juga sudah dapat diakses dimana saja pada zaman sekarang sehingga semua orang dapat menonton film kapanpun dan dimanapun. Film itu sendiri tidak hanya sedikit, jumlah film yang ada sangat banyak dan satu sama lain memiliki perbedaan baik dalam segi actor, cerita, dan direktur. Tapi hal yang paling mencolok dalam perbedaan suatu film adalah genre dari film tersebut. Genre merupakan hal utama yang memisahkan suatu kelompok film dengan kelompok film lainnya. Setiap genre memiliki topic atau focus berbeda dengan film dengan genre berbeda. Genre yang sama belum tentu memiliki plot cerita yang sama begitu juga sebaliknya, Plot yang mirip bisa saja memiliki genre yang berbeda. Menentukan genre yang tepat untuk suatu film berdasarkan plotnya memerlukan pengecekan berulang dengan bermacam-macam plot film yang memiliki genre yang sama. Pada penelitian ini film akan dapat diklasifikasikan terhadap genre yang ada hanya berdasarkan plot yang dimiliki dengan membandingkan plotnya dengan plot yang ada pada genre-genre yang tersedia, sehingga dapat ditemukan genre yang sesuai berdasarkan plot yang ada. Pembuatan program akan dilakukan dengan menggunakan Google Colaboratory dengan menggunakan bahasa python. Dalam penelitian ini akan dilakukan preprocessing yang prosesnya adalah data cleansing, case folding, stemming, stopwords removal, tokenizing, dan data reduction. Setelah melakukan preprocessing, klasifikasi akan dilakukan dengan menggunakan metode Naïve Bayes Classifier berdasarkan nilai bobot per kata yang diperoleh dengan menggunakan TF-IDF. Hasil dari klasifikasi akan didapatkan kata yang memiliki bobot terbesar dalam melakukan klasifikasi dan skor akurasi dari klasifikasi tersebut.

**Kata Kunci:** Plot, Film, Naïve Bayes Classifier, Tf-Idf, Accuracy Score, dan Python