

ABSTRAK

Saat ini perkembangan teknologi terus meningkat, berkat kemajuan ini lahirlah sebuah teknologi yang di namakan internet. Dengan adanya internet semua orang dapat dengan mudah terhubung dengan dunia luar. segala informasi dapat di akses dengan mudah, setiap harinya jutaan orang di Indonesia terhubung dengan internet. Salah satu mesin pencari yang paling populer di gunakan di dunia adalah *Google*, dengan milyaran pengguna setiap harinya. Salah satu *channel* milik *Google* yang populer bagi kalangan peneliti untuk mempublikasikan hasil penelitiannya adalah *Google Scholar*. *Google Scholar* dibuat untuk mempermudah pengguna mencari atau menelusuri materi maupun referensi pendidikan dan penelitian ilmiah dengan format publikasi. Referensi ini dapat berupa jurnal ilmiah yang merupakan hasil-hasil penelitian, riset atau kajian yang dapat di pertanggung jawabkan secara ilmiah.

Pada tugas akhir ini, dilakukan sebuah penelitian yang akan menggunakan metode *Analysis of Variance* (ANOVA). sebuah analisis statistik yang menguji perbedaan rata-rata data lebih dari dua kelompok. Fungsi dari algoritma ini sebagai alat analisis untuk menguji hipotesis penelitian yang mana menilai adakah perbedaan rerata antara kelompok. Data yang akan di kelompokkan dengan metode *Analysis of Variance* (ANOVA), sebelumnya akan di *crawling* terlebih dahulu dari *Google Scholar*.

Data dari *Google Scholar* berhasil di *crawling* dengan isi dokumen berbentuk pdf, lalu dokumen publikasi ilmiah tersebut diinputkan kedalam aplikasi R *shiny* untuk dianalisis lebih lanjut dengan menggunakan metode *Analysis of Variance* (ANOVA) untuk menghasilkan representasi kata kunci dari seluruh dokumen yang diuji. Dengan dibuatnya sistem ini, diharapkan dapat menjadi alat tes untuk mengetahui seberapa representatif *keyword* yang ada di dokumen ilmiah dengan keseluruhan isi dokumen.

Kata Kunci: *Analysis of Variance, data crawling, data mining, google scholar.*