# ABSTRAK

Penerimaan pegawai merupakan kunci utama dalam upaya meningkatkan kualitas sumber daya manusia di sebuah perusahaan. Baik atau buruknya pegawai sangat berpengaruh terhadap kualitas perusahaan. Perlu ketelitian dan waktu yang cukup lama dalam menyaring pelamar agar bisa mendapatkan calon pegawai yang berkompetensi, profesional dan sesuai dengan yang diharapkan. Akan tetapi sekarang masih banyak kantor cabang yang berada di kota-kota kecil tidak mempunyai tenaga profesional yang mampu melakukan seleksi penerimaan pegawai. Maka dari itu perlu adanya pendekatan data mining dalam melakukan seleksi penerimaan pegawai.

Data mining merupakan metode gabungan antara analisa tradisional dengan algoritma canggih yang berguna untuk menganalisa pola atau kecenderungan dari sekumpulan data. Dengan data mining diharapkan seleksi penerimaan pegawai bisa berlangsung lebih efektif, efisien dan harapannya mampu mengurangi tingkat subjektifitas. Pada penelitian ini algoritma *Classification and Regression Trees* diimplementasikan untuk melakukan pengklasifikasian data penerimaan pegawai.

Adapun datanya diambil dari penilaian penerimaan Tenaga Harian Lepas di Kantor Satuan Polisi Pamong Praja Kabupaten Pati pada tahun 2018. Data tersebut berjumlah 290 calon pegawai, 230 calon pegawai ditolak dan 60 calon pegawai diterima. Untuk membentuk pohon keputusan permasalahan ini menggunakan perhitungan indeks gini. Tahap awal pengolahan data adalah dengan cara membersihkan data dari atribut yang tidak terpakai terlebih dahulu. Selanjutnya data yang bertipe numerik ditransformasikan menjadi tipe nominal. Setelah itu data bisa diproses dengan menggunakan sistem klasifikasi penerimaan pegawai. Lalu hasil klasifikasi dari sistem diuji menggunakan *K-Fold Cross Validation*, dengan nilai K = 5 dan K = 10, apabila pengujian menggunakan jumlah fold yang lain maka pembagian datanya menjadi tidak rata dan hasil pengukurannya menjadi bias. Pengujian menggunakan 5-*Fold Cross Validation* menghasilkan tingkat akurasi sebesar 98,96%, presisi sebesar 99,13% dan *recall* sebesar 96,88%. Sedangkan pengujian menggunakan 10-*Fold Cross Validation* menghasilkan tingkat akurasi sebesar 88,96%, presisi sebesar 96,73% dan *recall* sebesar 97,33%. Dari perbandingan pengujian tersebut, pengujian menggunakan K = 5 menghasilkan nilai akurasi dan presisi yang lebih baik sedangkan pengujian menggunakan K = 10 menghasilkan nilai *recall* yang lebih baik.

Maka pengujian yang lebih optimal untuk data yang digunakan dalam penelitian ini adalah 5-*Fold Cross Validation*. Menurut data penerimaan Tenaga Harian Lepas di Kantor Satuan Polisi Pamong Praja Kabupaten Pati pada tahun 2018 ada 60 calon pegawai yang diterima dan 230 calon pegawai ditolak. Sedangkan dengan pengujian 5-*Fold Cross Validation* ada 59 calon pegawai yang diterima dan 231 calon pegawai yang ditolak. Perbedaan itu disebabkan karena adanya 3 data calon pegawai yang tidak terklasifikasi dengan benar. Untuk penelitian selanjutnya diharapkan melakukan pengujian menggunakan algoritma yang lain atau metode pengujian lainnya, sehingga dapat mengetahui perbandingan hasil dengan penelitian ini. Dan dari perbandingan tersebut diharapkan penelitian selanjutnya mampu memperbaiki kekurangan penelitian ini.

Kata Kunci: Klasifikasi, Penerimaan Pegawai, Classification and Regression Trees, K-Fold Cross Validation