

ABSTRAK

ANALISIS DAN DETEKSI *FRAUD* PADA DATA PANGGILAN MENGUNAKAN ALGORITMA *K-NEAREST NEIGHBOR* (STUDI KASUS: PT XYZ)

Oleh

EPTA RIZKY ALFIYYAH

NIM : 1202164077

Penelitian ini membahas mengenai salah satu permasalahan telekomunikasi yang terjadi di Indonesia yaitu *telecom fraud*. Salah satu Badan Usaha Milik Negara (BUMN) yaitu PT XYZ, merupakan pihak yang bertanggung jawab dalam menangani perihal *telecom fraud* tersebut. *Telecom fraud* merupakan suatu aktivitas penggunaan fasilitas telekomunikasi yang dilakukan secara ilegal dan disengaja dalam berbagai bentuk kecurangan, penipuan, atau pun penggelapan oleh orang atau organisasi tertentu yang tujuannya adalah mendapatkan layanan tersebut dan menghindari biaya layanan atau pelacakan rekaman tagihan yang dilakukan secara ilegal. Tujuan penelitian ini yaitu untuk menganalisis dan mendeteksi kecurangan (*fraud*) pada data panggilan yang terindikasi sebagai *SIMBox fraud* dimana ini telah merugikan pihak PT XYZ yang mempunyai tugas menangani masalah *fraud* tersebut. Penelitian ini dilakukan dengan menggunakan proses *data mining* dan algoritma *K-Nearest Neighbor*. Terdapat beberapa jenis *telecom fraud*, *SIMBox fraud* merupakan salah satu dari jenis *telecom fraud* yang akan dibahas pada penelitian ini. Data yang digunakan dalam penelitian ini merupakan data panggilan yang dilakukan pada bulan Juni hingga Agustus tahun 2017, dimana data ini nantinya akan diproses menggunakan algoritma *k-nearest neighbor*. Data panggilan ini berupa data csv yang diolah atau diproses dengan menggunakan proses *data mining*. *Data mining* adalah proses yang menggunakan teknik statistik, matematika, kecerdasan buatan, dan *machine learning* untuk mengekstraksi dan mengidentifikasi informasi yang bermanfaat dan pengetahuan yang terkait dari berbagai *database* besar. Hasil dari penelitian ini nantinya berupa tingkat akurasi

pengimplementasian algoritma *K-Nearest Neighbor* dalam mengklasifikasi data nomor telepon yang terindikasi sebagai *SIMBox fraud*. Penelitian dilakukan menggunakan *tool* Jupyter Notebook. Setelah dilakukan pengujian algoritma *K-Nearest Neighbor* dengan menggunakan $k=1$, $k=3$, $k=5$, $k=7$ dan $k=9$, maka diperoleh nilai akurasi tertinggi yaitu 74.2% dengan nilai $k=9$, rasio 0.1, *precision* 74%, *recall* 74%, dan *f1-score* 74%.

Kata kunci : *Telecom Fraud, SIMBox fraud, Data Mining, K-Nearest Neighbor.*