

Nama : Wahyu Adi Nugroho

NIM : A11.2019.12310

Kelp. : 46UG

Tugas Minggu Ke – 8 Data Mining

Carilah 3 paper tentang perkembangan aplikasi Data Mining menggunakan ANN dan Deep Learning minimal 5 tahun terakhir (terbit antara thn 2017 – 2022), kemudian review paper tersebut, selanjutnya tuliskan kedalam paper A4 minimal 1 halaman penuh.

Penerapan Deep Learning dalam Deteksi Penipuan Transaksi Keuangan Secara Elektronik			
Tahun	Penulis	Link	Algoritma
2021	Faried Zamachsari, Niken Puspitasari	http://jurnal.iaii.or.id/index.php/RESTI/article/view/2952/391	Machine Learning with Ensemble dan Deep Learning
Review			
<p>Tujuan penelitian ini adalah untuk mendapatkan model terbaik sehingga dapat disimulasikan ke dalam sistem pembayaran perbankan dengan tujuan melakukan deteksi penipuan transaksi keuangan secara elektronik sehingga dapat mencegah terjadinya kerugian kepada nasabah dan perbankan. Kedua jenis machine learning tersebut akan dikonfigurasi dan dioptimalisasi menggunakan SMOTE sehingga mendapatkan model dengan nilai accuracy score dan precision score terbaik. Data transaksi keuangan diambil dari simulasi pembayaran bank yang dibangun di atas konsep Multi Agent-Based Simulation (MABS) oleh bank di Spanyol dengan tujuan untuk mempromosikan pengembangan aplikasi Big Data. Penelitian sebelumnya menggunakan data yang sama dengan judul A Bayesian Classifier based on Constraints of Ordering of Variables for Fraud Detection mendapatkan accuracy score 99.272%. Untuk membangun model terbaik tidak hanya mempertimbangkan accuracy score akan tetapi precision score merupakan nilai yang perlu dipertimbangkan. Precision score sangat menentukan untuk pencegahan terjadinya penipuan. Deteksi penipuan mendapatkan hasil terbaik tanpa proses SMOTE dengan menggunakan Deep Learning dengan accuracy score yaitu 99.602% dan precision score 90.574%. Dengan melakukan penambahan SMOTE akan meningkatkan accuracy score dan precision score dengan model terbaik yang dihasilkan pada Extra Trees Classification dengan nilai accuracy score 99.835% dan precision score 99.786%.</p>			

Nama : Wahyu Adi Nugroho

NIM : A11.2019.12310

Kelp. : 46UG

Analisis Data Mining Untuk Memprediksi Lama Perawatan Pasien Covid-19 Di DIY			
Tahun	Penulis	Link	Algoritma
2022	Agung Supoyo, Putri Taqwa Prasetyaningrum	https://ejournal.bsi.ac.id/ ejurnal/index.php/ Bianglala/article/view/11890	Random Forest, k-NN dan Deep Learning
Review			
<p>Penelitian ini menggunakan dataset yang diperoleh dari Dinas Kominfo DIY untuk kasus periode Maret sampai dengan September 2020. Diperlukan preprocessing (data reduction, data cleaning dan data integration) sebelum dilakukan analisis data mining. Preprocessing menghasilkan dataset sejumlah 271 record data dengan 31 kolom. Analisis data mining menggunakan algoritma Random Forest, k-NN dan Deep Learning menghasilkan performance model dengan RMSE masing-masing sebesar 4,949; 6,349 dan 5,436. Setelah dilakukan seleksi atribut untuk optimalisasi dihasilkan nilai RMSE sebesar 4.817 pada algoritma Random Forest dengan menggunakan 23 atribut. Hasil analisis belum cukup baik jika dibandingkan dengan rata-rata lama perawatan sebesar 15.339 hari karena menghasilkan NRMSE sebesar 31,40%. Nilai performance model dipengaruhi oleh pemilihan atribut yang digunakan. Lima atribut yang paling berpengaruh terhadap lama perawatan pasien adalah usia, jenis kelamin, kecamatan, batuk. Untuk meningkatkan performance model diperlukan penelitian lanjutan menggunakan record data yang lebih banyak dengan tambahan atribut lain seperti rumah sakit perawatan dan tindakan medis.</p>			

Penerapan Data Mining Untuk Klasifikasi Penyakit Hepatocellular Carcinoma Menggunakan Algoritma Naïve Bayes			
Tahun	Penulis	Link	Algoritma
2021	Bambang Tri Rahmat Doni, Sari Susanti, Ade Mubarok	http://ejurnal.ars.ac.id/index.php/ jti/article/view/403	Naïve Bayes
Review			

Nama : Wahyu Adi Nugroho

NIM : A11.2019.12310

Kelp. : 46UG

Tujuan penelitian ini adalah untuk mengklasifikasikan tingkat kemungkinan hidup pasien yang telah di diagnosis menderita penyakit Hepatocellular Carcinoma dengan menggunakan penerapan metode data mining serta melakukan pengukuran terhadap performa algoritma Naïve Bayes dengan mengacu kepada Confusion Matrix dan Kurva ROC. Data sekunder yang digunakan merupakan data publik yang bernama HCC Survival Data Set yang dirilis pada tahun 2017 dan diperoleh dari website UCI Machine Learning Repository. Algoritma Naïve Bayes merupakan salah satu algoritma yang terdapat dalam metode data mining yang menerapkan Teori Keputusan Bayes pada teknik klasifikasi dengan menggunakan cara pendekatan statistik yang bersifat fundamental dalam pengenalan pola. Teknik validasi yang digunakan menggunakan teknik 10-Fold Cross-Validation dengan rasio pembagian data sebesar 90:10. Perangkat lunak yang digunakan adalah RapidMiner Studio v9.5. Hasil penelitian menunjukkan bahwa hasil performa algoritma Naïve Bayes yang diukur menggunakan Confusion Matrix dengan nilai yang dihasilkan berupa nilai Akurasi sebesar 70,30%, Presisi sebesar 73,53% dan Recall sebesar 77,32% serta hasil performa yang diukur menggunakan Kurva ROC (Receiver Operating Characteristic) dengan nilai yang dihasilkan berupa nilai AUC sebesar 0.783 yang termasuk dalam kategori Fair Classification atau kategori Klasifikasi Cukup.