**Nama : Welly Oktaviandi**

**NIM : A11.2018.11026**

**Mk : SA\_Data Mining**

Jurnal “Classification of Breast Cancer Patients Using Neural Network Technique”

Pada jurnal ini dilakukan data mining dengan menggunakan Teknik artificial neural network untuk mengklasifikasikam pasien kanker payudara. Artificial neural network adalah suatu metode pengolahan informasi yang memiliki karakteristik yang sama dengan jaringan saraftiruanbiologis.Jaringansarafinitelahdikembangkanuntukmenyimpulakankesimpulan secara umum dengan model matematika berdasarkan pemikiran manusia. Jurnal ini mengklasifikasikan kasus kangker ini menjadi kelas ganas dan kelas jinak yang menunjukkan jika suatu orang kangker. Aplikasi ini dikembangkan dengan menggunakan software Matlab yang dapat mengklasifikasikan pasien kanker secara cepat dan akurat. Aplikasi ini bertujuan untuk mengembangkan suatu aplikasi yang dapat mengklasifikasikan pasien kanker payudara serta melakukan klasifikasi pasien kanker payudara. Untuk pekerjaan ini, data pasien kanker payudaradiperolehdariWilliamH.WolbergdiRumahSakitUniversitasWisconsin,Madison, Wisconsin, AS. Dataset ini memiliki 699 kelas / kasus. Ini diunduh dari situs web Machine LearningRepository.

Pada jurnal ini dijelaskan bahwa untuk mengolah data ini mereka menggunakan metodelogi CRISP-DM. penggunaan CRISP-DM ini dilakukan karena dapat mengidentifikasi suatumasalahdansystemdapatdikembangkansesuaidengankriteriadankebutuhanuserserta kebutuhan system. ANN memiliki 2 arsitektur yaitu single level arsitektur dan multi level arsitektur

Jurnal ini dilakukan agar menghasilkan klasifikasi yang cepat dan akurat dengan menggunakanmodelANN.ModelANNinimenghasilkanrata-rataklasifikasisebesar98.88% dan dengan kecepatan 0.182 detik. (Badarudin *et al.*,2021)

Jurnal“ComparisonofMachineLearningv/sDeepLearningModeltoPredictICD9code using Text MiningTechniques”

Pada jurnal ini dilakukan perbandingan antara machine learning dengan deep learning untuk memprediksikan ICD9 menggunakan teks mining. Pada jurnal ini terdapat permasalahan berupamengidentifikasidanmembandingkanTeknikvektorisasiterbaik,yaituTF,IDFdanTF- IDF. Hal ini dilakukan dengan tujuan untuk mengetahui apakah deep learning dapat mengungguli machinelearning.

Untuk melakukan perbandingan ini, maka dilakukan dengan metode CRISP-DM pada model Deep learning. Menggunakan metode ini karena memberikan cara yang terorganisir dan metode yang kuat serta all around.

CRISP-DM yang dilakukan adalah

* Business Understanding: Memahami tujuan dan prasyarat proyek dan menentukan tujuan dengan datamining.
* Data Understanding: bermacam-macam informasi awal dan aklimatisasi dibuat dan menggunakan teknik EDA untuk mengidentifikasi masalah kualitasinformasi.
* Data Preparation: Penentuan data dan sifat serta pembersihandata.
* Modeling: Memisahkan dan menguji data antar model, membandingkan model yang berbeda.
* Evaluation: Menentukan apakah hasil dari tahap pemodelan memenuhi persyaratan bisnis
* Deployment: Menerapkan proses ke penggunaan waktunyata.

Pada jurnal ini didapatkan kesimpulan berupa TF-IDF adalah Teknik vektorisasi terbaik bila dibandingkan dengan TF dan IDF. Model deep learning menghasilkan performa yang kuat sebelum cross validation, namun lemah setelahnya. Karena itu deep learning tidak lebih baik daripada machine learning. (Bhat, 2020)

Jurnal “Deep Learning Versus Traditional Solutions for Group Trajectory Outliers”

Pada jurnal ini dilakukan perbandingan antara deep learning dengan solusi tradisional untuk The group trajectory outlier. Pada jurnal ini diberikan beberapa solusi untuk menangani group of trajectory outliers, yaitu data mining, computational Inteligence, dan deep learning.

Pada solusi data mining diberikan usulan algoritma DBSCAN-GTO,KNN-GTO dan FS-GTO

Pada solusi computational intelligence diberikan usulan algoritma Genetic Algorithm dan Particel Swarm Optimization

Pada solusi deep learning diberikan usulan algoritma berupa CNN-GTO yang menggunakan convolutional neural network untuk mengidentifikasi group of trajectory outliers. Terdapat beberapa tahap pada CNN-GTO, yaitu pengumpulan data, training dan Inference:.

Pada pengumpulan data dilakukan dengan tujuan mengumpulkan data dan membentuk image database (gambaran database). Strategi visual digunakan, di mana setiap gambar berisi serangkaian lintasan perilaku manusia. Prosesnya dimulai dengan merekam frame video dari kamera. frame tersebut kemudian diubah menjadi gambar, dan teknik distorsi yang berbeda seperti pemetaan, pengubahan ukuran, digunakan untuk mengoreksi gambar. Seluruh gambar lintasan disimpan ke dalam database untuk training

Pada training dilakukan pengubahan group of trajectory outlier detection menjadi masalah deteksi objek dengan mengatur input model deteksi objek ke database gambar lintasan, dan output model deteksi objek ke group of trajectory outliers di setiap gambar. Untuk setiap gambardiinput,wilayahminatditentukandanditeruskankelapisantersembunyitempatfungsi aktivasi Relu dilakukan. Proses serupa dengan jaringan saraf konvolusi diikuti. Dalam RCNN cepat, model berperforma lebih baik dan lebih cepat karena wilayah minat ditemukan menggunakan metode penelusuran selektif, dan semua wilayah minat gambar ditemukan sekaligus. Ini berbeda dengan CNN yang menemukan ROI (wilayah minat) dan menerapkan Relu pada setiap ROI secara terpisah, yang lebih lambat. Proses ini diulangi untuk sejumlah periode tertentu, atau hingga pelatihan berhenti memberikan peningkatan untuk sejumlah iterasi tertentu. Inisialisasi bobot dilakukan dengan menggunakan model ImageNet yang telah dilatih sebelumnya. Model terlatih disimpan di centralworkstation.

Inference dilakukan dengan tujuan untuk mendapatkan input gambar group of trajectory outliersmenggunakanmodelterlatihdarilangkahsebelumnya.Denganbegitupropagasibobot yang berbeda dari model terlatih dilakukan untuk mendeteksi objek gambar. Objek yang terdeteksidianggapsebagaigroupoftrajectoryoutliers.Padalangkahini,berbagaijenis

inferensi dihasilkan. Penulis kemudian mengirim model yang dilatih ke komputer dan menyimpulkanmodeluntuksetiapdatagambarlintasanbaru.Penulisjugadapatmenggunakan smartphone yang mendukung Andorid, dan komputasi GPU untuk menyimpulkan kelompok pencilanlintasandalampemrosesanwaktunyata.Dalamkonteksini,beberapateknologidapat diintegrasikan seperti TensorflowLite 2. (Belhadi *et al.*,2020)

Badarudin, P. M. *et al.* (2021) ‘Classification of Breast Cancer Patients Using Neural Network Technique’, 1, pp. 13–19.

Belhadi, A. *et al.* (2020) ‘Deep Learning Versus Traditional Solutions for Group Trajectory Outliers’, *IEEE Transactions on Cybernetics*, pp. 1–12. doi: 10.1109/TCYB.2020.3029338.

Bhat, A. (2020) ‘Comparison of Machine Learning v / s Deep Learning Model to Predict ICD9 code using Text Mining Techniques’.