

Penerapan Algoritma *Resticted Boltzmann Machine* Pada Pemilihan Bidang Minat Mahasiswa Informatika Universitas Muhammadiyah Malang

Bella Chintia Eka Merita^{*1}, Zamah Sari², Vinna Rahmayanti Setyaning Nastiti³, Briansyah Setio Wiyono

^{1,2,3}Universitas Muhammadiyah Malang

bella.merita1@gmail.com^{*1}, zamahsari@umm.ac.id², vinastiti@umm.ac.id³, brian@umm.ac.id

Abstrak

Pemilihan bidang minat masih menjadi sebuah kesulitan untuk sebagian mahasiswa, terutama mahasiswa yang belum percaya diri dengan kemampuan dan keahliannya. Bidang minat pada jenjang perguruan tinggi menjadi sebuah tolak ukur dan kunci kesuksesan di masa mendatang sebagai penunjang karirnya. Sehingga mahasiswa harus memiliki pemikiran yang matang dan tepat sasaran agar dapat mencapai tujuannya. Untuk menyelesaikan masalah ini dibutuhkan suatu hasil pembelajaran sebagai bahan pertimbangan dalam menentukan bidang minat mahasiswa. Dengan hasil pembelajaran tersebut mahasiswa dapat mempertimbangkan pilihannya berdasarkan grafik nilai dan bidang minat yang telah diambil oleh mahasiswa yang sudah lulus terlebih dahulu. Dengan demikian dilakukan penelitian untuk memberikan hasil pembelajaran dalam bentuk klasifikasi bidang minat mahasiswa Informatika Universitas Muhammadiyah Malang menggunakan data mahasiswa angkatan 2013 – 2015 yang sudah lulus. Data tersebut dilakukan pengklasifikasian menggunakan algoritma *Restricted Boltzmann Machine*, dan dilakukan pengujian model menggunakan *confusion matrix*. Data penelitian yang digunakan merupakan data yang memiliki kelas tidak seimbang karena jumlah masing-masing bidang minat berbeda jauh. Sehingga dilakukan pengklasifikasian menggunakan *SMOTE* untuk menangani ketidakseimbangan kelas. Hasil akurasi yang didapatkan pada *RBM+SMOTE* yaitu sebesar 70% dan nilai *mean square error* 0.4.

Kata Kunci: Klasifikasi, Peminatan, Deep Learning, *Restricted Boltzmann Machine*, *SMOTE*

Abstract

Choosing an interest field/area is still a difficulty for some students, especially students who are not confident in their abilities and expertise. The interest field/area at the tertiary level is a benchmark and the key for the future to success as a support for his career. So the students need to think carefully on the target to achieve their target. Solving the problem of a learning result is needed as a consideration in determining the student's interest field/area. Along with these learning outcomes, the students can consider their choices based on the graph of the values and a the interest field/area that have been taken by students who have graduated first. Thus, a study is conducted to provide learning outcomes in the form of classifications of the interest field/ area of Informatics students at the University of Muhammadiyah Malang using data from students class 2013 - 2015 who had graduated. The data is classified using the *Restricted Boltzmann Machine* algorithm, and the model is tested using *confusion matrix*. The accuracy results obtained are 66%, but these results are still not good enough because the prediction results refer only to one of the three interest fields/areas. Then the test is done using *RBM* and *SMOTE* by balancing the minority and majority classes. From the implementation of the combination of the two algorithms, an accuracy of 70% is obtained.

Keywords: Classification, Specialization, Deep Learning, *Restricted Boltzmann Machine*, *SMOTE*

1. Pendahuluan

Dalam dunia pendidikan jenjang perguruan tinggi pemilihan bidang minat studi merupakan hal yang penting pada proses belajar mahasiswa. Bidang minat bertujuan mengarahkan peserta didik agar lebih fokus mengembangkan kemampuan diri dan dan minat yang dimiliki [1]. Bagi mahasiswa, pemilihan bidang minat yang tepat akan berdampak pada proses belajar juga dalam

menyelesaikan tugas akhir agar dapat lulus tepat waktu. Dalam penelitian psikologi pendidikan, minat bakat ikut serta memberikan pengaruh terhadap prestasi akademik [2].

Universitas perlu memanfaatkan secara optimal baik dari fasilitas, mahasiswa dan pendidik, infrastruktur, pengetahuan dan data yang dimiliki [3]. Data dapat diproses dan diolah sedemikian rupa menjadi sebuah informasi yang dapat membantu berbagai masalah, seperti pada pemilihan bidang minat mahasiswa. Dalam pemilihan bidang minat data yang dapat dimanfaatkan yaitu riwayat nilai dan bidang minat yang telah ditempuh mahasiswa yang sudah lulus. Dengan data tersebut dapat dilihat dari grafik nilai dan sebagai acuan pertimbangan pemilihan bidang minat. Dalam membantu menentukan bidang minat dapat dipertimbangkan berdasarkan beberapa faktor. Faktor internal merupakan faktor yang berasal dari dalam diri individu, sedangkan faktor eksternal merupakan faktor yang berasal dari luar diri individu [4].

Pada sebagian mahasiswa masih memiliki kesulitan dalam menentukan pilihan yang tepat. Hal tersebut dikarenakan mahasiswa belum sepenuhnya mengetahui kemampuan dan keahlian di suatu bidang tertentu. Lain halnya dengan mahasiswa yang sudah memiliki nilai hasil studi baik dan kemantapan pola pikir menjadi suatu hal yang mudah dalam menentukan pilihan bidang minat. Ketepatan pemilihan bidang minat mahasiswa juga dapat berpengaruh terhadap akreditasi program studi, sehingga kesulitan yang sedang dialami mahasiswa menjadi bahan evaluasi penyelenggara akademik agar dapat membantu mahasiswa lulus tepat waktu

Menangani masalah kesulitan pemilihan bidang minat mahasiswa, sebelumnya telah dilakukan penelitian ilmiah yang bertujuan membantu memecahkan masalah yang sedang dihadapi. Seperti pada penelitian [5], menentukan jurusan pada SMAN 16 Semarang menggunakan metode *K-Nearest Neighbor* dengan mengklasifikasikan data siswa berdasarkan nilai rapor, nilai ujian, minat penjurusan, rekomendasi BK, dan hasil penjurusan. Pada penelitian [6] menerapkan metode *Forward Selection* pada algoritma *Support Vector Machine* dan *Naïve Bayes Classifier Kernel Density* pada klasifikasi jalur minat SMA.

Berdasarkan penelitian [7] menerapkan algoritma fuzzy naïve bayes dalam pemilihan bidang keahlian Mahasiswa Informatika Universitas Islam Lamongan berdasarkan nama, hobby, mata kuliah favorit, IPK dan bidang minat. Metode *Invariant Moment* dan *Restricted Boltzmann Machine* [8] digunakan untuk pengenalan tulisan tangan dengan menggunakan data hasil tanda tangan 10 koresponden dengan jumlah sebanyak 150 data. *Restricted Boltzmann Machine* juga digunakan pada masalah klasifikasi dengan membuat tipe baru dalam struktur organisasinya [9].

Ditinjau dari permasalahan pada kesulitan mahasiswa menentukan pemilihan bidang minat dan pemanfaatan data mahasiswa, maka akan dilakukan pengklasifikasian data mahasiswa Informatika untuk memberikan pertimbangan dalam menentukan bidang minat menerapkan metode *deep learning* [10] dengan menggunakan algoritma *Restricted Boltzmann Machine*. Dengan penelitian ini diharapkan dapat melihat hasil dari kinerja algoritma *Restricted Boltzmann Machine* pada klasifikasi data mahasiswa Informatika UMM untuk pemilihan bidang minat. Selain itu juga diharapkan dapat membantu pihak penyelenggara akademik Informatika UMM menghasilkan mahasiswa yang berkompeten dan memiliki keahlian sesuai dengan bidangnya

2. Metode Penelitian

2.1 Bidang Minat

Bidang minat merupakan sebuah wadah untuk menampung seseorang sesuai dengan kemampuannya pada pekerjaan tertentu. Minat berpengaruh dalam aktivitas seseorang pada kehidupan sehari-hari, serta dapat berpengaruh pada keputusan-keputusan penting dalam hidup seseorang [11]. Pada program studi Informatika Universitas Muhammadiyah Malang terdapat tiga bidang studi yaitu Rekayasa Perangkat Lunak, Jaringan, dan Game Cerdas. Dengan adanya bidang minat tersebut diharapkan dapat mencetuskan lulusan mahasiswa yang berkualitas dan memiliki kompetensi dalam bidang tertentu.

2.2 Pengumpulan Data

Pada Gambar 1, pengumpulan data yang digunakan yaitu data mahasiswa Informatika 2013-2015 dengan melakukan perizinan pengambilan data di kantor jurusan Informatika. Data yang telah dikumpulkan yaitu data mahasiswa dan nilai hasil studi dari semester satu hingga masa studi yang ditempuh selesai. Variabel penelitian yang digunakan meliputi nilai mata kuliah, Pemrograman Terstruktur, Pengantar Teknologi Informasi, Logika Informatika, Interaksi Manusia & Komputer, Organisasi dan Arsitektur Komputer, Pemrograman Objek dengan Java, Pengantar Topologi, Algoritma dan Struktur Data, Grafika Komputer I, Desain Database, Komunikasi Data,

Gambar 1. Data Penelitian

2.3 Preprocessing Data

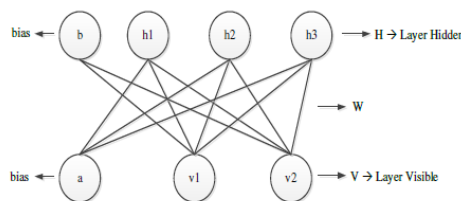
Preprocessing data merupakan langkah yang dilakukan sebelum dataset diinputkan dalam model. Tujuan dilakukan preprocessing data yaitu untuk membersihkan data, mengisi nilai yang *missing value*, menangani *inconsistent* data dan data *noise*. Berikut macam-macam tahapan *preprocessing* data.

Data cleaning, pada tahapan ini dilakukan untuk membersihkan data *noise*, menangani data *noise*, menangani *inconsistent* data dan menangani *outlier*. *Data integration*, tahapan ini dilakukan untuk menggabungkan data dari beberapa *database* atau beberapa penyimpanan data. Tahapan ini bertujuan untuk mengurangi dan menghindari redundansi dan *inconsistent*.

Data reduction, tahap *reduction* dilakukan untuk mengurangi data dari data yang volume besar menjadi lebih kecil namun tetap menjaga integritas data asli. *Data transformation*, pada tahapan ini data diubah kedalam bentuk sesuai yang dibutuhkan agar hasil yang didapatkan lebih efisien. Normalisasi diterapkan pada tahap ini, dengan mengubah data atribut menjadi skala tau berada dalam range 0,0 hingga 1,0.

2.4 Restricted Boltzmann Machine

Pada tahun 2006, (Hinton et al) menyediakan cara yang efektif untuk pembelajaran yang mendalam yang dinamakan *Deep Belief Network*. Pada DBN, *Restricted Boltzmann Machine* untuk melatih bobot jaringan lapis demi lapis. RBM merupakan model berbasis energi juga dikatan sebagai jenis dari *Markov Random Field* (MRF) [12]. RBM terdiri dari lapisan terlihat dan tidak terlihat atau *hidden layer* dan *visible layer*. Neuron tersembunyi mewakili vector fitur untuk diletakkan pola. Jumlah neuron tersembunyi memiliki peran peting dalam klasifikasi [9]. Arsitektur RBM dijelaskan pada Gambar 2 dengan Persamaan 1 berikut.



Gambar 2. Arsitektur RBM

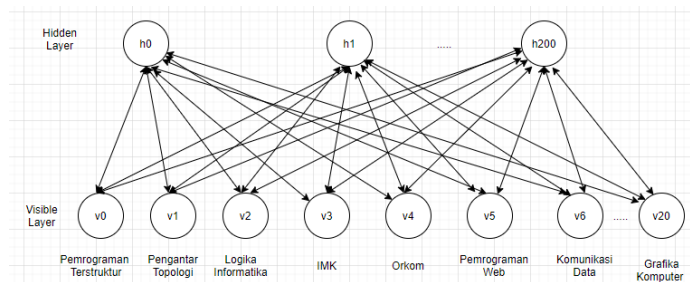
$$E(V, H) = - \sum_i a_i V_i - \sum_j b_j H_j - \sum_{i,j} V_i H_j W_{i,j} \tag{1}$$

Algoritma tersebut dapat diterapkan pada permasalahan data dimensi tinggi seperti gambar dan teks, serta digunakan pada masalah klasifikasi, pemodelan tekstur, pemodelan

gerak, pengambilan informasi dan pengenalan. Selain digunakan untuk klasifikasi pada penelitian [10] [11], juga telah digunakan pada permasalahan rekomendasi sistem seperti pada rekomendasi film berdasarkan rating yang disukai user [13], dan prediksi potensi efek samping dari narkoba menggunakan rekomendasi sistem pada [14]. Dalam permasalahan *clustering* dapat dilihat pada penelitian mengenai pengelompokan pembicara dan pelacakan tugas di acara TV [15]. Berdasarkan beberapa penelitian yang telah dilakukan dalam penerapan RBM, maka dapat dikatakan bahwa algoritma tersebut dapat diimplementasikan pada berbagai masalah pada algoritma pembelajaran dengan pengawasan dan tanpa pengawasan.

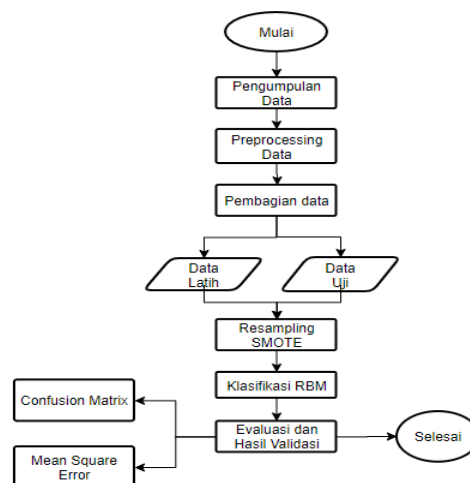
2.5 Skema Jaringan RBM

Gambar 3 merupakan langkah pertama dalam melatih RBM dengan banyak input. Input dikalikan dengan bobot dan kemudian ditambahkan ke bias, hasilnya kemudian dilewatkan melalui fungsi aktivasi sigmoid dan output menentukan keadaan hidden nodes diaktifkan atau tidak. Bobot akan menjadi matriks dengan jumlah node input sejumlah baris dan jumlah node tersembunyi sebagai jumlah kolom. Node tersembunyi pertama akan menerima perkalian vector dari input dikalikan dengan kolom bobot pertama sebelum bias yang sesuai ditambahkan. Pada pemilihan bidang minat nilai x yang digunakan yaitu nilai setiap mata kuliah sebagai input untuk mendapatkan nilai sample dari pelatihan yang telah didistribusikan. Pada pelatihan RBM tidak ada *visible node* dan *hidden node* yang saling terhubung ke *visible node* lain atau *hidden node* lainnya. Untuk melatih RBM, sampel dari training set yang digunakan sebagai masukan untuk RBM melalui neuron visible, dan kemudian jaringan bolak-balik antara neuron visible dan hidden. Tujuan dari pelatihan adalah untuk pembelajaran koneksi bobot pada visible atau hidden dan bias aktivasi neuron sehingga RBM belajar untuk merokonstruksi data input selama fase dimana sampel neuron visible dari neuron hidden. Sampling antara hidden dan visible dikontrol oleh learning rate. Pada library sklearn Bernoulli RBM fungsi yang digunakan yaitu fungsi aktivasi sigmoid yang sudah secara default terdapat pada library tersebut.



Gambar 3. Skema Jaringan RBM pada pemilihan bidang minat

2.6 Perancangan Implementasi dan Pengujiannya



Gambar 4. Skema implementasi algoritma dan pengujiannya

2.7 Random Oversampling

Class imbalance atau ketidakseimbangan kelas pada sebuah dataset merupakan suatu permasalahan dalam *machine learning*. Hal tersebut dikarenakan kelas minoritas jauh lebih kecil daripada kelas mayoritas. Ketidakseimbangan data akan menghasilkan akurasi prediksi pada kelas minoritas yang rendah [16]. Selain itu juga dapat mempengaruhi kualitas data dalam kinerja klasifikasi. Permasalahan ketidakseimbangan data dapat diatasi dengan pendekatan teknik *resampling* dimana teknik ini dilakukan untuk menyeimbangkan data asli dengan menyesuaikan jumlah sample pada kelas berbeda. Terdapat tiga kategori dalam pendekatan *resampling* yaitu *over-sampling*, *under-sampling*, dan *hibrida*. *Over-sampling* merupakan teknik dimana kelas minoritas diduplikasi agar dapat menyeimbangi jumlah kelas mayoritas. Salah satu teknik populer pada teknik *over-sampling* yaitu (*Synthetic Minority Over-sampling Technique*). Pada awalnya, pendekatan metode SMOTE tidak meningkatkan kinerja klasifikasi karena duplikasi data yang diberikan semakin meningkatkan pola minoritas. Namun hasil pada [17] dapat dilihat bahwa dengan menggunakan metode SMOTE menunjukkan pengaruh yang signifikan pada masalah klasifikasi

2.8 Pengukuran Tingkat Akurasi Model

Pengujian model klasifikasi dilakukan dengan menggunakan *confusion matrix* dan pengukuran *mean square error*. Hasil prediksi model dapat dilihat dari tabel *confusion matrix* [18] dengan membandingkan data aktual dan data prediksi, sehingga dapat diketahui jumlah prediksi salah dan prediksi benar serta hasil klasifikasi setiap kelasnya. Evaluasi hasil klasifikasi menggunakan *mean square error* [19]. Jika hasil akurasi dirasa sudah baik maka proses klasifikasi sudah selesai, namun jika dirasa hasil kurang baik maka dilakukan analisa terhadap data dengan mencari akar permasalahan dan melakukan percobaan pengujian menggunakan teknik lain.

Table 1. Confusion Matrix

		Actual	
		Yes	No
Prediction	Yes	TP	FP
	No	FN	TN

Berikut keterangan Tabel 1:

- TP (*True Positive*) → Jumlah data dengan nilai sebenarnya positif dan nilai prediksi positif
- FP (*False Positive*) → Jumlah data dengan nilai sebenarnya negatif dan nilai prediksi positif
- FN (*False Negative*) → Jumlah data dengan nilai sebenarnya positif dan nilai prediksi negatif
- TN (*True Negative*) → Jumlah data dengan nilai sebenarnya negatif dan nilai prediksi negative

Berikut Persamaan 2, Persamaan 3, dan Persamaan 4 untuk menghitung nilai akurasi, *precision*, dan *recall* berdasarkan tabel *Confusion Matrix* yang telah dibuat sebelumnya.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

2.9 Mean Square Error

Mean Squared Error merupakan metode untuk mengevaluasi metode peramalan [19]. Metode ini merupakan rata-rata kuadrat antara nilai yang diramalkan dan yang diamati. Kesalahan kecil yang dihasilkan memiliki kemungkinan menghasilkan perbedaan yang besar. Persamaan 5 berikut rumus menghitung MSE sebagai berikut.

$$MSE = \sum_t^n = 1 \frac{(X_t - F_t)^2}{n} \quad (5)$$

3. Hasil dan Pembahasan

3.1 Analisa Pengumpulan Data

Pada tahap ini pengumpulan data yang digunakan yaitu data mahasiswa Informatika 2013-2015 dengan melakukan perizinan pengambilan data di kantor jurusan Informatika. Dari data berjumlah 800 dilakukan pengurangan data yang diperlukan sehingga mendapatkan jumlah 319 data mahasiswa.

3.2 Pembagian Data

Pembagian data sebanyak 80% data training dan 20% data test. Jumlah keseluruhan data sebelum dilakukan resampling sebanyak 319. Setelah dilakukan resampling data menjadi 588. Sehingga didapatkan jumlah data training sebanyak 470, dan data testing sebanyak 118 data.

3.3 Implementasi Algoritma RBM dan Teknik SMOTE

Pada implementasi algoritma RBM menggunakan Bahasa pemrograman python dan dengan menggunakan perangkat lunak anaconda navigator. Gambar 5 dan Gambar 6 berikut serta dapat dilihat langkah-langkah membangun model RBM:

1. Memasukkan dataset yang akan digunakan
2. Melakukan processing data dengan menghilangkan missing value dan mengisi nilai tersebut
3. Mengubah data dari nilai matakuliah, bidang minat dan lama masa studi menjadi bilangan biner
4. Menginisialisai nilai x dan y
5. Melakukan resampling data menggunakan SMOTE
6. Menginisialisai nilai training dan testing
7. Pembangunan model RBM

```

panj = Counter(y)
# define oversampling strategy SMOTE
strategy = {0:196, 1:196, 2:196}
oversample = SMOTE(sampling_strategy=strategy)
# fit and apply the transform SMOTE
X_over, y_over = oversample.fit_resample(X, y)
panj_over = Counter(y_over)
scaler = StandardScaler()
X_over = scaler.fit_transform(X_over)

```

Gambar 5. Source code membangun model RBM

```

#RBM
logistic = linear_model.LogisticRegression(solver='newton-cg', tol=1)
logistic.C = 6000
rbm = BernoulliRBM(random_state = 28, verbose=False, learning_rate=0.000001, n_iter=1000, n_components=500)
rbm_features_classifier = Pipeline(steps=[('rbm', rbm), ('logistic', logistic)])
rbm_features_classifier.fit(X_train, y_train)

```

Gambar 6. Source code membangun model RBM

Dataset yang digunakan pada penelitian memiliki data yang tidak seimbang (*imbalanced data*), maka dilakukan pengujian *balancing* data agar kelas mayoritas tidak mendominasi kelas minoritas dengan menyeimbangkan data menggunakan teknik *oversampling* dan dikombinasikan dengan algoritma SMOTE. Jumlah data game pada awalnya jauh lebih sedikit dibandingkan jaringan dan RPL. Jumlah data pada RPL mendominasi yaitu sejumlah 196 data, Jaringan 102 dan jumlah Game 20. Disinilah terlihat jelas perbedaan yang sangat jauh antara game dan RPL. Dengan menggunakan teknik *oversampling* jumlah data minoritas akan diduplikasi sesuai jumlah kelas mayoritas.

3.4 Implementasi Naïve Bayes

Penerapan algoritma naïve bayes pada penelitian ini, dilakukan untuk mengetahui hasil perbandingan antara RBM dengan metode lainnya. apakah hasil yang didapatkan lebih baik RBM atau lebih baik menggunakan metode lainnya. Metode naïve bayes merupakan salah satu algoritma pada teknik klasifikasi. Naïve bayes diasumsikan bahwa ada atau tidak ada ciri tertentu dari sebuah kelas tidak ada hubungannya dengan ciri dari kelas lainnya [20].

3.5 Evaluasi Model Klasifikasi

Dalam penelitian ini evaluasi model klasifikasi menggunakan Confusion Matrix untuk mengukur nilai akurasi model yang didapatkan. untuk mengukur nilai error maka dilakukan pengukuran menggunakan mean square error. Hasil akurasi model RBM dapat dilihat pada Tabel 2.

Table 2. Hasil akurasi model RBM

	Pengujian	Precision	Recall	Akurasi	MSE
1	RBM	0.43	0.66	0.66	0.34
2	RBM+SMOTE	0.68	0.70	0.70	0.4
3	Naïve Bayes	0.60	0.58	0.57	0.5
4	Naïve Bayes + SMOTE	0.65	0.66	0.65	0.5

Dari tabel dapat diketahui bahwa hasil akurasi menggunakan model RBM dan menggunakan teknik resampling data SMOTE menghasilkan akurasi sebesar 70%. Dapat dikatakan bahwa hasil ini cukup baik untuk digunakan pada data yang tidak seimbang.

4. Kesimpulan

Berdasarkan penelitian yang telah dilakukan mengenai klasifikasi bidang minat menggunakan *Restricted Boltzmann Machine* didapatkan hasil sebagai berikut:

1. Dataset yang digunakan memiliki kelas data yang tidak seimbang atau *imbalancing class*
2. Menggunakan teknik oversampling yaitu SMOTE untuk menangani masalah *imbalancing class*
3. Hasil akurasi RBM lebih baik dibandingkan dengan Naïve bayes
4. Hasil akurasi RBM + SMOTE lebih baik dibandingkan naïve bayes + SMOTE
5. Membuktikan bahwa RBM digunakan dalam klasifikasi pemilihan bidang minat mahasiswa Informatika UMM memiliki hasil yang baik jika menggunakan SMOTE pada dataset yang memiliki kelas tidak seimbang.

Daftar Notasi

Contoh penulisan notasi dapat diuraikan dengan keterangan sebagai berikut:

- i : mewakili jumlah visible neuron
- j : mewakili jumlah hidde neuron
- V_i : binary state dari visible unit i
- H_i : binary state dari hidden unit j
- b_i : bias dari hidden unit
- W_{ij} : bobo tantara visible dan hidden unit

Referensi

- [1] Y. S. Nugroho, "Klasifikasi dan Klastering Penjurusan Siswa SMA Negeri 3 Boyolali," *Khazanah Inform. J. Ilmu Komput. dan Inform.*, vol. 1, no. 1, p. 1, 2015.
- [2] N. Fartindyah and Subiyanto, "Sistem pendukung keputusan peminatan sma menggunakan metode," *J. Pendidik.*, vol. 44, no. november 2014, pp. 139–145, 2013.
- [3] E. Miranda, "Implementasi Data Warehouse dan Data Mining: Studi Kasus Analisis Peminatan Studi Siswa," *ComTech Comput. Math. Eng. Appl.*, vol. 2, no. 1, p. 1, 2011.
- [4] P. Sosial, D. I. Lingkungan, and S. Dan, "Studi Tentang Hubungan Dukungan Sosial, Penyesuaian Sosial Di Lingkungan Sekolah Dan Prestasi Akademik Siswa Smpit Assyfa Boarding School Subang Jawa Barat," *Stud. Tentang Hub. Dukungan Sos. Penyesuaian*

- Sos. Di Lingkungan. Sekol. Dan Prestasi Akad. Siswa Smpit Assyfa Board. Sch. Subang Jawa Barat*, vol. 10, no. 2, pp. 103–114, 2011.
- [5] A. Sulistiyo, "Penentuan Jurusan Sekolah Menengah Atas Menggunakan Metode K-Nearest Neighbor Classifier Pada SMAN 16 Semarang," *Fasilkom Udinus*, vol. 1, no. 1, pp. 1–5, 2014.
- [6] T. B. Sasongko and O. Arifin, "Implementasi Metode Forward Selection pada Algoritma SVM dan Naive Bayes Classifier," vol. 6, no. 4, pp. 383–388, 2019.
- [7] N. Fuad, "Pemanfaatan Algoritma Fuzzy Naive Bayes Dalam Pemilihan Bidang Keahlian Mahasiswa Teknik Informatika Universitas Islam Lamongan," *J. Tek.*, vol. 11, no. 2, p. 1117, 2019.
- [8] H. A. Majid, K. E. Dewi, and J. D. Bandung, "Signature Recognition Using Invariant Moment Method And Restricted Boltzmann Machine."
- [9] S. Pirmoradi, M. Teshnehlab, N. Zarghami, and A. Sharifi, "The Self-Organizing Restricted Boltzmann Machine for Deep Representation with the Application on Classification Problems," *Expert Syst. Appl.*, vol. 149, p. 113286, 2020.
- [10] J. Wira and G. Putra, "Pengenalan Pembelajaran Mesin dan Deep Learning Jan Wira Gotama Putra Pengenalan Konsep Pembelajaran Mesin dan Deep Learning," no. March 2018, 2019.
- [11] F. A. Nurcahyo, S. Azwar, and W. Martani, "Stimulus Gambar: Sebuah Kajian pada Instrumen Minat Vokasional," *Bul. Psikol.*, vol. 26, no. 2, p. 111, 2018.
- [12] N. Zhang, S. Ding, J. Zhang, and Y. Xue, "An overview on Restricted Boltzmann Machines," *Neurocomputing*, vol. 275, pp. 1186–1199, 2018.
- [13] A. Pujahari and D. S. Sisodia, "Modeling Side Information in Preference Relation based Restricted Boltzmann Machine for recommender systems," *Inf. Sci. (Ny)*, vol. 490, pp. 126–145, 2019.
- [14] W. Zhang, H. Zou, L. Luo, Q. Liu, W. Wu, and W. Xiao, "Predicting potential side effects of drugs by recommender methods and ensemble learning," *Neurocomputing*, vol. 173, pp. 979–987, 2016.
- [15] J. Hernando, "applied sciences Restricted Boltzmann Machine Vectors for Speaker Clustering and Tracking Tasks in TV," 2019.
- [16] O. Heranova, "Synthetic Minority Oversampling Technique pada Averaged One Dependence Estimators untuk Klasifikasi Credit Scoring," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 3, no. 3, pp. 443–450, 2019.
- [17] D. Elreedy and A. F. Atiya, "A Comprehensive Analysis of Synthetic Minority Oversampling Technique (SMOTE) for handling class imbalance," *Inf. Sci. (Ny)*, vol. 505, pp. 32–64, 2019.
- [18] L. Swastina, "Penerapan Algoritma C4.5 untuk Penentuan Jurusan Mahasiswa," *J. Gema Aktual.*, vol. 2, no. 1, pp. 93–98, 2013.
- [19] U. Prediksi, P. Pada, P. Tertentu, K. M. S, and S. P. W, "Metode Single Exponential Smoothing Untuk Prediksi Penjualan Pada Periode Tertentu (Studi Kasus: PT. Media Cemara Kreasi) 1," no. 1998, pp. 259–266, 2015.
- [20] H. Naparin, "Klasifikasi Peminatan Siswa SMA Menggunakan Metode Naive Bayes," *Syst. Inf. Syst. Informatics J.*, vol. 2, no. 1, pp. 25–32, 2016.