



PREDIKSI HASIL UJIAN NASIONAL SISWA MENENGAH ATAS : PENDEKATAN DATA MINING

Putra Prima¹, Aris Fanani², Billy Montolalu³, Ummiy Fauziah Laili⁴

¹Politeknik Negeri Malang, ²UIN Sunan Ampel Surabaya, ³Universitas Bhayangkara
Surabaya, ⁴IAIN Kediri

putraprima@gmail.com¹, arisfa@uinsby.ac.id², billy.montolalu@gmail.com³,
ummiyfauziyah@gmail.com⁴

Abstrak: Pemerintah Indonesia menggunakan Nilai Ujian Nasional (NUN) digunakan sebagai standar nasional untuk mengukur keberhasilan siswa. Dengan menggunakan NUN sebagai parameter untuk mengukur tingkat keberhasilan siswa akan ada resiko yang sangat besar jika siswa gagal dalam menjalani Ujian Nasional (UN), oleh karena itu dibutuhkan sebuah model untuk melakukan prediksi terhadap NUN siswa sehingga dapat dilakukan usaha pencegahan terhadap gagalnya siswa dalam UN. Pada paper ini diusulkan sebuah metode untuk melakukan prediksi terhadap NUN siswa dengan menggunakan neuro fuzzy. Data yang digunakan dalam penelitian ini adalah nilai rapor siswa, dan nilai IQ siswa. Pada penelitian ini digunakan tiga mata pelajaran sebagai subjek penelitian yaitu Matematika, Bahasa Indonesia dan Bahasa Inggris. Berdasarkan hasil pengujian yang dilakukan menunjukkan bahwa metode neuro fuzzy dengan menggunakan data nilai dan data IQ memberikan akurasi prediksi terbaik pada mata pelajaran matematika sebesar 74%.

Kata Kunci: prediksi, nilai UN, neuro fuzzy

Abstract: The Indonesian government uses the National Exam Score (NUN) as a national standard to measure student success. By using NUN as a parameter to measure the level of student success there will be a very big risk if students fail to take the National Examination (UN), therefore a model is needed to predict student NUN so that efforts can be made to prevent student failures in the UN. This paper proposes a method to predict student NUN using neuro fuzzy. The data used in this study were student report cards and student IQ scores. In this study, three subjects were used as research subjects, namely Mathematics, Indonesian and English. Based on the results of the tests conducted, it showed that the neuro-fuzzy method using value data and IQ data provided the best predictive accuracy in mathematics at 74%.

Keywords: prediction, UN value, neuro fuzzy

Pendahuluan

Kemampuan memprediksi secara akurat menjadi suatu kebutuhan pada saat ini. Ketidاكلulusan siswa mengikuti Ujian Nasional (UN) menjadi permasalahan rutin setiap tahun bagi siswa, orang tua, sekolah dan pemerintah. Banyak metode prediksi menggunakan pendekatan data miningdikembangkan selama beberapa tahun terakhir. Prediksi awal keberhasilan siswa dengan menggunakan data pendaftaran siswa (Kovacic, 2010) menunjukkan bahwa faktor yang penting memisahkan siswa yang berhasil dari siswa yang gagal adalah etnis dan program bimbingan belajar.

Pada penelitian (Marquez-Vera, 2011) kegagalan siswa dapat ditentukan dengan menggunakan data dari performa siswa dan data socio demografi. (Casanova,2005) juga menyimpulkan bahwa kondisi dari keluarga siswa juga memiliki pengaruh terhadap pada prestasi siswa. Pada penelitian yang lain (Rosyidi, 2011) mengusulkan prediksi dari kegagalan siswa menggunakan data nilai raport

siswa. Sayangnya metode yang diusulkan hanya menggunakan data nilai raport siswa, padahal ada beberapa faktor yang dapat mempengaruhi prediksi, seperti data socio demografi dan lingkungan belajar.

Berdasarkan penelitian yang sebelumnya dilakukan, pada penelitian ini diusulkan metode neuro-fuzzy untuk memprediksi hasil Ujian Nasional (UN) Siswa Menengah Atas (SMA) tidak hanya menggunakan data nilai raport tetapi juga menggunakan data IQ. Dengan menambahkan data IQ, akan didapatkan peningkatan akurasi dari prediksi nilai UN dengan menggunakan neuro-fuzzy.

Metodologi Penelitian

Penelitian ini mengambil data dari SMA di Kota Lamongan. Data diambil dengan cara survey langsung dan diambil data siswa angkatan tahun 2018/2019 dengan mata pelajaran yang digunakan adalah Bahasa Indonesia, Bahasa Inggris dan Matematika.

Pembahasan

Ujian Nasional

Ujian Nasional (UN) adalah kegiatan pengukuran dan penilaian kompetensi peserta didik secara nasional pada jenjang pendidikan dasar dan menengah. Sedangkan Nilai UN (NUN) adalah nilai yang diperoleh oleh peserta didik dalam mengikuti UN.

UN dilaksanakan oleh BNSP bekerjasama dengan instansi terkait di lingkungan pemerintah, pemerintah propinsi, pemerintah kabupaten/kota, perguruan tinggi dan satuan pendidikan. Untuk tingkat Sekolah Menengah Atas (SMA), ada 3 mata pelajaran utama yang digunakan, yaitu : Matematika, Bahasa Indonesia, dan Bahasa Inggris.

Neuro Fuzzy

Neuro-fuzzy adalah penggabungan mekanisme *fuzzy inference system* yang digambarkan dalam arsitektur jaringan syaraf. Sistem inferensi fuzzy yang digunakan adalah sistem inferensi fuzzy model Tagaki-Sugeno-Kang (TSK) orde satu dengan pertimbangan

kesederhaan dan kemudahan komputasi (Jang, 1997).

Pada sistem logika fuzzy memiliki kemampuan lebih dalam menangani data pengetahuan lingkungan luar serta keupayaan dalam persepsi dan penalaran seperti otak manusia. Namun demikian sistem logika fuzzy tidak memiliki kemampuan untuk belajar dan beradaptasi. Sebaliknya pada sistem jaringan syaraf tiruan memiliki kemampuan untuk belajar dan beradaptasi, akan tetapi tidak memiliki kemampuan penalaran seperti yang dimiliki pada sistem logika fuzzy. Untuk mendapatkan sebuah kecerdasan yang memiliki kemampuan penalaran serta kemampuan pembelajaran, sehingga berkembang Neuro Fuzzy.

Integrasi antara sistem fuzzy dengan jaringan syaraf memiliki beberapa karakteristik, antara lain (Jang, 1997) : (a) mampu mengimplementasikan kepakaran manusia, (b) model komputasi diinspirasi secara biologis, (c) menggunakan komputasi numeris, (d) mendukung domain untuk aplikasi-aplikasi baru, seperti pemrosesan sinyal

secara adaptif, identifikasi sistem non linear, regresi non-linear dan pencocokan pola, (e) pembelajaran dilakukan dengan model yang tidak terikat, (f) proses komputasi dilakukan secara intensif, (g) memiliki toleransi terhadap kegagalan, (h) dapat digunakan untuk aplikasi-aplikasi dunia nyata.

1.1 Struktur Neuro Fuzzy

Struktur Neuro Fuzzy yang menggambarkan sistem fuzzy TSK dapat digambarkan dalam diagram blok atau arsitektur jaringan syaraf seperti pada Gambar 1. Himpunan aturan umum untuk Sugeno orde satu dengan dua input adalah sebagai berikut:

Aturan 1: if x is A_1 and y is B_1 , then $f_1 = p_1x + q_1y + r_1$

Aturan 2: if x is A_2 and y is B_2 , then $f_2 = p_2x + q_2y + r_2$

Layer 1. Setiap node i pada lapisan ini adalah node adaptif dengan sebuah fungsi node,

$$\mu(x) = \frac{1}{1 + \left| \frac{x - c_i}{a_i} \right|^{2b_i}} \quad (1)$$

dimana $\{a_i, b_i, c_i\}$ merupakan himpunan parameter.

Layer 2. Setiap node pada lapisan ini merupakan node yang berisi bobot sudah tetap, dimana adalah hasil dari semua sinyal yang masuk.

$$O_{2,i} = \mu A_i(x) * \mu B_i(y), \quad i = 1, 2 \quad (2)$$

Setiap node putput merepresentasikan kekuatan mengirim dari aturan.

Layer 3. Setiap node pada lapisan ini adalah node tetap. Node ini menghitung perbandingan kekuatan mengirim aturan ke- i dengan total dari kekuatan mengirim aturan secara keseluruhan.

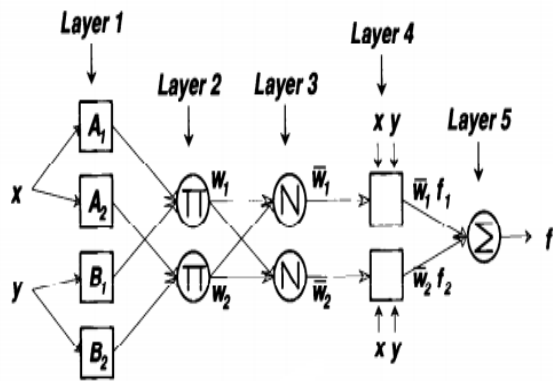
$$O_{3,1} = \bar{w} = \frac{w_i}{w_1 + w_2}, \quad i = 1, 2 \quad (3)$$

Output pada lapisan ini disebut *normalized firing strengths*.

Layer 4. Setiap node pada lapisan ini adalah node adaptif dengan fungsi node, $O_{4,1} = \bar{w}_i f_i = \bar{w}_i(p_i x + q_i y + r_i)$, (4) dimana \bar{w}_i adalah kekuatan mengirim yang sudah dinormalisasikan pada lapisan ke-3 dan $\{p_i, q_i, r_i\}$ merupakan himpunan parameter.

Layer 5. Node tunggal pada lapisan ini adalah node tetap, dimana memperhitungkan keseluruhan output sebagai hasil dari sinyal yang masuk.

$$O_{5,1} = \sum \bar{w}_i f_i = \frac{\sum_i \bar{w}_i f_i}{\sum_i w_i} \quad (5)$$



Gambar 1. Arsitektur Neuro Fuzzy dengan 2 Input dan 1 Output

Tabel 1. Pembelajaran hybrid

Parameter	Arah Maju	Arah Mundur
Parameter Premis	Tetap	Gradient Distance
Parameter Konsekuen	RLSE	Tetap

1.2 Pembelajaran pada Neuro Fuzzy

Struktur Neuro Fuzzy seperti pada Gambar 1, simpul adaptif terdapat pada lapisan pertama dan keempat. Simpul pada lapisan pertama mengandung parameter premis non-linear, sedangkan pada lapisan keempat mengandung parameter linier. Untuk memperbaharui parameter-parameter tersebut dilakukan proses pembelajaran dengan algoritma hybrid. Pembelajaran hybrid terdiri atas dua bagian yaitu arah maju (*forward Pass*) dan arah mundur (*backward pass*).

Pada arah maju, parameter premis dibuat tetap dengan menggunakan metode *Recursive Least Squar Estimator* (RLSE). Pada RLSE parameter konsekuen diperbaiki berdasarkan pasangan data masukan dan keluaran. Kemudian setelah parameter konsekuen didapatkan, data masukan dilewatkan jaringan adaptif kembali dan hasil keluaran jaringan adaptif ini dibandingkan dengan keluaran yang sebenarnya.

Pada arah mundur, parameter konsekuen dibuat tetap. Kesalahan yang terjadi antara keluaran jaringan adaptif dan keluaran sebenarnya dipropagasikan balik dengan menggunakan *gradient distance* untuk memperbaiki parameter premis. pembelajaran ini disebut *backpropagation-error*. Satu tahap arah pembelajaran maju-mundur dinamakan satu epoch. Pembelajaran hybrid neuro fuzzy seperti dijelaskan pada Tabel 1.

Nilai nilai untuk masing masing mata pelajaran dan hasil tes IQ siswa digunakan sebagai masukan untuk sistem neuro fuzzy yang digunakan, dengan demikian terbentuk sistem neuro

fuzzy yang mempunyai enam masukan dan satu nilai keluaran.

Sebelum dilakukan proses prediksi, dataset nilai dan IQ dibersihkan dahulu agar nilai prediksi menjadi lebih akurat. Pembersihan data ini meliputi pencarian titik tengah, penghapusan *outlier*, pencarian data kembar dan normalisasi. Pencarian data yang kembar dilakukan dengan menggunakan query. Penghapusan *outlier* dilakukan dengan cara mencari titik tengah kemudian dijadikan acuan untuk mendapatkan nilai minimum dan maksimum yang diijinkan.

Pengolahan data dilakukan dengan menggunakan tool anfisedit

pada matlab, dengan jumlah data sebanyak 108. Data dibagi menjadi dua bagian yaitu data training dan data testing. Untuk data training digunakan 60% dari data, sedangkan untuk data testing digunakan 40% dari data. Contoh tampilan data (*training* dan *testing*) seperti Tabel 2.

Data ditraining dengan menggunakan tool anfisedit pada matlab 2010. Fungsi keanggotaan yang digunakan adalah trimf, gbellmf dan gaussmf, fungsi keanggotaan ini digunakan dalam proses training untuk mengetahui mana fungsi keanggotaan yang paling cocok untuk digunakan terhadap dataset yang ada.

Tabel 2. Contoh Data (*Training* dan *testing*)

IQ	Smt1	Smt2	Smt3	Smt4	Smt5	NUN	IQ	Smt1	Smt2	Smt3	Smt4	Smt5	NUN
95	65	70	75	70	69	7.4	100	75	68	70	70	83	8
104	75	68	70	65	71	8.8	98	65	68	75	70	76	7
97	75	68	70	70	79	6.8	94	70	68	75	70	79	7.2
103	65	68	80	70	72	9.6	101	65	69	80	70	72	9.6
106	70	68	80	70	75	8.4	94	70	70	70	70	70	8.8
101	65	69	70	70	70	8.8	103	65	68	80	70	72	9.6
94	70	68	75	70	73	7.8	97	75	68	70	70	79	6.8
94	70	70	70	70	70	7.8	95	75	68	75	70	73	7.8
95	75	68	75	70	77	7.8							

Tabel 3. Rata-Rata Error Hasil Prediksi

Mata Pelajaran	Trimf (%)	Gbellmf (%)	Gaussmf (%)
Bahasa Indonesia	27,24	30.60	29.16
Bahasa Inggris	38,74	35,91	33,98
Matematika	26,90	32,16	32.56

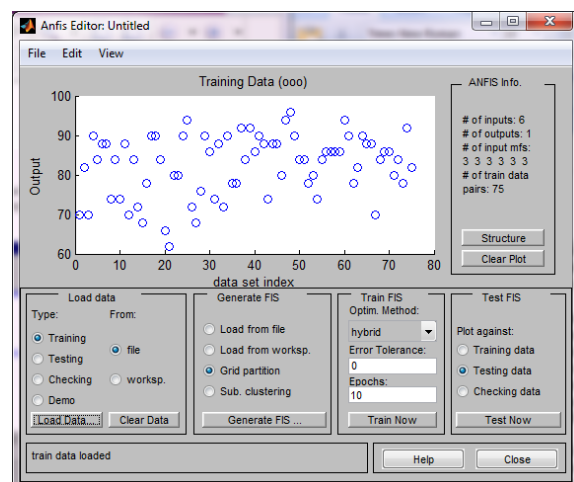
Sebagai contoh pengujian terhadap nilai Bahasa Indonesia, pengolahan diawali dengan memasukkan input data yang telah dinormalisasi dengan editor seperti Gambar 2, dan struktur neuro fuzzy seperti Gambar 3. Dengan menggunakan fungsi keanggotaan Trimf dengan *epoch* 10 dan memasukkan data testing didapatkan hasil pengujian seperti pada gambar 5. Hasil dari pengujian terhadap dataset dengan ketiga fungsi keanggotaan secara lengkap seperti ditampilkan pada Tabel 3.

Kesimpulan

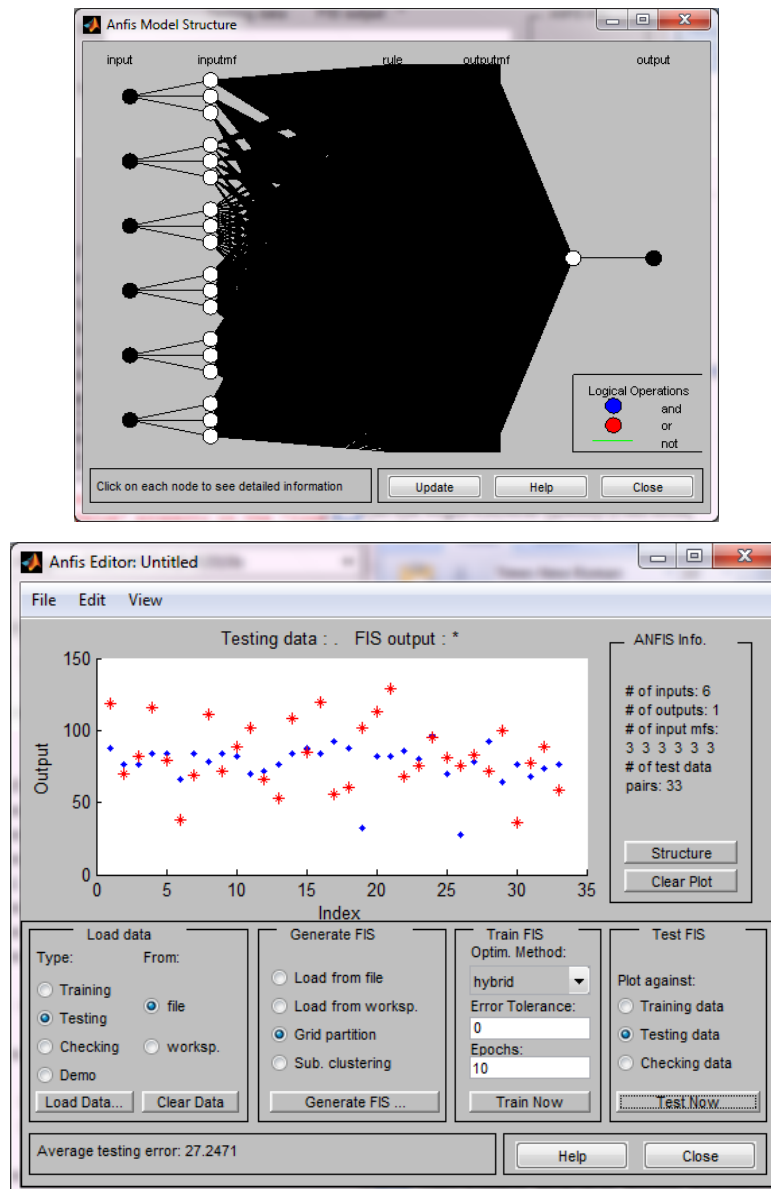
Nilai error diambil dari selisih antara nilai prediksi neuro fuzzy dengan nilai sesungguhnya. Dari tabel 2 dapat dilihat bahwa fungsi keanggotaan yang memiliki nilai error paling baik adalah fungsi keanggotaan trimf.

Diantara mata pelajaran yang di training mata pelajaran matematika mempunyai nilai error yang paling kecil sebesar 26,90% sedangkan yang memiliki nilai error terbesar adalah bahasa inggris sebesar 38,74%.

Error ini disebabkan karena terjadi perbedaan yang sangat mencolok antara nilai semester dengan nilai UN. Ini dapat terjadi karena standar penilaian seorang guru terhadap siswa dapat berbeda beda, terjadinya perubahan kurikulum, tingkat kesulitan soal UN yang berbeda beda dan kecurangan kecurangan yang sering terjadi pada saat dilaksanakannya ujian nasional.



Gambar 2. Input Data training



Gambar 3. Struktur Neuro Fuzzy dengan fungsi keanggotaan 3

Daftar Pustaka

Cortez, Paolo., Alice Silva., (2007), "Using Data Mining To Predict Secondary School Student Performance".

Kovačić, Z. J., (2010), "Early Prediction of Student Success: Mining Students Enrolment Data", Proceedings of Informing Science & IT Education Conference (InSITE).

Marquez-Vera, C., Romero, C., Ventura S., (2011), "Predicting School Failure Using Data Mining", 4th International Conference on Educational Data Mining, EDM.

Rosyidi, Imron., Moch Hariadi., I Ketut Eddy Purnama., (2011), "Data Mining Kemampuan Siswa Berbasis Neuro Fuzzy". SESINDO.

Shing Jyh, Roger Jang., (1997), "Neuro Fuzzy and Soft Computing", Prentice Hall