

## PREDIKSI KESIAPAN SEKOLAH MENGGUNAKAN *MACHINE LEARNING* BERBASIS KOMBINASI *ADAM* DAN *NESTEROV MOMENTUM*

Indah Mustika Rahayu<sup>\*1</sup>, Ahmad Yusuf<sup>2</sup>, Mujib Ridwan<sup>3</sup>

<sup>1,2,3</sup>Universitas Islam Negeri Sunan Ampel, Surabaya

Email: <sup>1</sup>indahrahayumustika@gmail.com, <sup>2</sup>ahmadyusuf@uinsby.ac.id, <sup>3</sup>mujibrw@uinsby.ac.id

<sup>\*</sup>Penulis Korespondensi

(Naskah masuk: 20 Agustus 2022, diterima untuk diterbitkan: 16 Desember 2022)

### Abstrak

Kesiapan sekolah adalah aspek perkembangan anak yang berperan pada kemampuan anak untuk beradaptasi dalam sistematisasi pendidikan tingkat dasar. Berdasarkan Permendikbud, usia 7 tahun adalah usia yang tepat bagi anak masuk Sekolah Dasar, karena anak telah memiliki kesiapan fisik dan psikis untuk mengikuti proses pendidikan formal. Namun, setiap anak tidak memiliki kondisi yang sama pada usia tertentu. Sehingga, diperlukan *Nijmeegse Schoolbekwaamheids Test (NST)* untuk mengukur kesiapan sekolah. Instrumen *NST* hanya dapat digunakan oleh Biro Psikologi yang mempunyai kemampuan dalam melakukan asesmen psikologis. Sedangkan, guru serta orang tua yang memiliki peran dalam bentuk pemberian dukungan dan stimulasi pada anak tidak dapat menggunakan instrumen tersebut. *Machine learning* adalah teknik yang menggunakan algoritma untuk menemukan pola yang berguna dalam data. Berdasarkan data *NST* terdahulu, dapat dirancang model prediksi kesiapan sekolah yang akan memudahkan guru dan orang tua dalam mengetahui kesiapan anak untuk masuk Sekolah Dasar. Data penelitian adalah data administratif 225 siswa yang telah mengikuti tes kesiapan sekolah berbasis *NST* yang diselenggarakan oleh TK Ar-Rasyid pada tahun 2012-2018. Data administratif *NST* terdiri dari umur, jenis kelamin, urutan anak, jumlah saudara, status TK, pendidikan ayah, pendidikan ibu dan hasil kesiapan sekolah. Berdasarkan korelasi *Chi-Square*, variabel yang memiliki hubungan signifikan kuat terhadap hasil tes kesiapan sekolah adalah status TK, jumlah saudara dan umur anak dengan nilai  $p < .05$ . Penelitian menggunakan algoritma *Artificial Neural Network* dengan metode optimasi kombinasi *Adam* dan *Nesterov Momentum*. Pengujian menggunakan skenario *5-fold cross validation*. Hasil penelitian menunjukkan bahwa kombinasi *Adam* dan *Nesterov Momentum* memiliki kinerja lebih baik daripada *classical Adam* dalam memprediksi kesiapan sekolah dengan akurasi 96% dan *loss* 0.06 dalam 1.98 *seconds* pada 10 *neuron* dan 1000 *epochs*.

**Kata kunci:** *machine learning, artificial neural network, adam optimizer, nesterov momentum, kesiapan sekolah*

## *SCHOOL READINESS PREDICTION USING MACHINE LEARNING BASED ON COMBINATION OF ADAM AND NESTEROV MOMENTUM*

### Abstract

*School readiness is an aspect of child development that plays a role in the child's ability to adapt in the systematics of elementary level education. Based on the Minister of Education and Culture, 7 years is the right age for children to enter elementary school, because children already have physical and psychological readiness to take part in the formal education process. However, every child does not have the same condition at a certain age. Thus, the Nijmeegse Schoolbekwaamheids Test (NST) is needed to measure school readiness. The NST instrument can only be used by the Psychology Bureau who has the ability to carry out psychological assessments. Meanwhile, teachers and parents who have a role in providing support and stimulation to children cannot use these instrument. Machine learning is a technique that uses algorithms to find useful patterns in data. Based on previous NST data, it can be designed as a school readiness prediction model that will facilitate teachers and parents in knowing the readiness of children to enter elementary school. Research data is administrative data of 225 students who have taken the NST-based school readiness test conducted by TK Ar-Rasyid in 2012-2018. NST administrative data consists of age, gender, child position, number of siblings, pre-elementary status, father education, mother education and school readiness results. Based on the Chi-Square correlation, variables that have a strong significant relationship to school readiness test results are pre-elementary status, number of siblings and age with  $p < .05$ . The research used Artificial Neural Network algorithms with a combination of Adam and Nesterov Momentum optimization method. Model testing used a 5-fold cross validation scenario. The results showed that the combination of Adam*

and Nesterov Momentum performed better than classical Adam in predicting school readiness with 96% accuracy and 0.06 loss in 1.98 seconds on 10 neurons and 1000 epochs.

**Keywords:** machine learning, artificial neural network, adam optimizer, nesterov momentum, school readiness

## 1. PENDAHULUAN

Kesiapan sekolah adalah aspek perkembangan anak yang meliputi kapasitas kognitif dan bahasa, perkembangan sosial-emosional, munculnya literasi dan berhitung, serta pengaturan diri yang berperan pada kemampuan anak untuk beradaptasi dan juga berkembang dalam sistematika sekolah tingkat dasar (Shaw, Mendelsohn, & Morris, 2021). Berdasarkan Peraturan Mendikbud Nomor 1 Tahun 2021 Pasal 2, pelaksanaan PPDB SD memprioritaskan penerimaan calon peserta didik baru kelas 1 SD yang berusia 7 tahun atau paling rendah 6 tahun bagi calon peserta didik yang memiliki kecerdasan atau bakat istimewa dan kesiapan psikis (Mendikbud, 2021).

Usia tujuh tahun dinilai tepat bagi anak untuk masuk sekolah dasar, karena pada usia tersebut anak telah mempunyai kesiapan fisik dan kesiapan psikis untuk mengikuti proses pendidikan Sekolah Dasar (Nurlina, Intan, & Syahriani, 2019). Sedangkan, usia kurang dari 7 tahun dinilai kurang tepat, karena anak belum mampu untuk mengikuti sistematika sekolah. Ketidaksiapan anak untuk mengikuti peraturan dan sistematika sekolah dapat berdampak pada gangguan komunikasi dan pengendalian emosi pada usia remaja sampai usia dewasa (Wulansuci & Kurniati, 2019).

Namun, setiap anak tidak memiliki kondisi fisik dan psikis yang sama pada usia tertentu (Wardhani, Jaya, & Fauziyah, 2019). Jenis kelamin juga dapat memengaruhi perbedaan masa perkembangan anak (Herdyana, 2019). Sehingga, diperlukan instrumen *Nijmeegse Schoolbekwaamheids Test (NST)* untuk memperkirakan kesiapan anak memasuki pendidikan Sekolah Dasar yang meliputi kesiapan fisik dan juga psikis. Tujuan *NST* adalah untuk mendeteksi siswa Taman Kanak-kanak yang barangkali menghadapi kesulitan dalam mengikuti proses awal pendidikan Sekolah Dasar (Susilarini, 2021).

Namun, instrumen *NST* hanya dapat digunakan oleh Biro Psikologi yang mempunyai kemampuan dalam melakukan assessment psikologis. Sehingga, guru serta orang tua yang memiliki peran dalam bentuk pemberian dukungan dan stimulasi pada anak tidak dapat menggunakan instrumen tersebut (Fitri & Reza, 2020). *Machine Learning* adalah teknik yang menggunakan algoritma untuk menemukan pola yang berguna dalam data. Berdasarkan data *NST* terdahulu, dapat dilakukan pemodelan prediksi kesiapan sekolah yang akan memudahkan guru dan orang tua mengerti kesiapan anak untuk masuk Sekolah Dasar.

Algoritma yang telah digunakan dalam prediksi kesiapan sekolah adalah *Decision Tree*, *K-Nearest Neighbor* dan *Artificial Neural Network* (Nurlina et al., 2019) (Intan, Ghani, & Salman, 2019). *Artificial Neural Network (ANN)* dengan metode *optimization*

*Stochastic Gradient Descent* menghasilkan *accuracy* lebih baik dibandingkan dengan algoritma *Decision Tree* dan juga *K-Nearest Neighbor* dalam melakukan klasifikasi kesiapan sekolah.

*Artificial Neural Network* dengan *Stochastic Gradient Descent* telah memiliki *performance* yang baik dalam berbagai aplikasi (Cheridito, Jentzen, & Rossmannek, 2021). Namun, metode optimasi *SGD* menetapkan skala gradien secara seragam ke segala arah, sehingga hal tersebut dapat merugikan kinerja *ANN* jika skala *gradient* bernilai kecil (Wang, Zhou, & Zhong, 2018). Untuk memperbaiki kekurangan tersebut, metode adaptif telah diusulkan yang secara diagonal menetapkan skala gradien dengan perkiraan nilai kelengkungan atau kurva fungsi *error*. Metode tersebut adalah *Adaptive Moment Estimation* atau dikenal dengan *Adam* (Kingman & Ba, 2015).

*Adam* mempunyai kinerja lebih baik daripada metode *SGD* (Guna & Putri, 2020). Metode *Adam* juga memiliki kemampuan yang lebih cepat dalam proses *training* dibandingkan dengan metode *SGD* (Huda, Mubarak, & Adiwijaya, 2019). Komponen *Adam* terdiri dari momentum dan *adaptive learning rate*. Komponen momentum yang digunakan oleh metode *Adam* berupa *vanilla momentum*. Sehingga, pada segi momentum, metode *Nesterov Momentum* memiliki kinerja lebih baik daripada *Adam*.

*Adam* dan metode *Nesterov Momentum* dapat dikombinasi untuk meningkatkan kinerja *Adam* (Dozat, 2016). *Nesterov Momentum* adalah metode untuk memperbaiki *vanilla momentum*. Dalam kasus perubahan parameter *learning rate*, momentum dapat mengalami penurunan kecepatan dan peningkatan *loss* karena osilasi yang besar. Untuk menghindari osilasi besar, gradien akan diarahkan pada parameter sebelumnya. Hal tersebut yang dilakukan *Nesterov Momentum* atau *Nesterov Accelerated Gradient*.

Berdasarkan riset sebelumnya, dapat diketahui bahwa *Adam* menghasilkan kinerja yang baik dalam *ANN* dengan menetapkan skala gradien berdasarkan perkiraan kelengkungan fungsi *error*. *Nesterovs* juga dapat menghasilkan nilai *loss* yang minimum dalam *ANN* dengan menentukan gradien ke arah parameter sebelumnya untuk menghindari osilasi. Sehingga, kombinasi *Adam* dan *Nesterov* dapat dilakukan untuk mempercepat proses *training* data, mengurangi nilai *loss* dan meningkatkan nilai akurasi prediksi.

Penelitian ini bertujuan untuk menguji kinerja metode kombinasi *Adam* dan *Nesterovs Momentum* pada *Artificial Neural Network*. Metode kombinasi ini diharapkan mampu meningkatkan hasil prediksi kesiapan sekolah. Sehingga, hasil penelitian ini dapat dimanfaatkan guru dan orang tua untuk mengetahui kesiapan anak untuk masuk sekolah tingkat dasar.

## 2. METODE PENELITIAN

### 2.1. Data Understanding

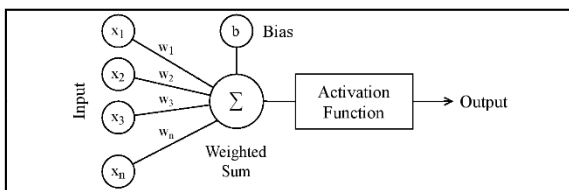
Data penelitian berasal dari data administratif setiap anak yang telah mengikuti tes kesiapan masuk sekolah tingkat dasar berbasis instrumen *Nijmeegse Schoolbekwaamheid Test* yang diselenggarakan oleh TK Ar-Rasyid pada tahun 2012-2018 dengan diikuti 225 siswa. Data administratif terdiri dari umur, jenis kelamin, urutan anak, jumlah saudara, status TK, pendidikan ayah, pendidikan ibu dan hasil kesiapan sekolah. Deskripsi variabel ditunjukkan pada Tabel 1.

Tabel 1. Deskripsi Variabel (Nurlina et al., 2019)

Attribute	Description	Value
Age	Age	int
Gender	Gender	0: Laki-laki 1: Perempuan
position	Child position	int
siblings	Number of siblings	int
preschool	Have pre-elementary	0: Tidak 1: Ya
fatheredu	Father education	0: SMP 1: SMA 2: Diploma 3: Sarjana
motheredu	Mother education	0: SMP 1: SMA 2: Diploma 3: Sarjana
readiness (label)	School readiness	0: Tidak Disarankan 1: Dipertimbangkan 2: Disarankan

### 2.2. Modeling

*Training data* menggunakan algoritma *artificial neural network* dengan beberapa skenario parameter untuk menemukan *architecture ANN* dengan kinerja terbaik. *Artificial neural network* tersederhana terdiri dari satu neuron dan disebut *perceptron*. Perhitungan *activation function* pada sebuah neuron ditunjukkan pada diagram berikut.



Gambar 1. Single Perceptron Illustration (Guna & Putri, 2020)

*Activation function* didapatkan dari jumlah dari perkalian nilai *input* dan *weight*. *Weight* merupakan nilai acak yang telah diinisialisasi. Penentuan nilai *weighted sum* menggunakan persamaan (1).

$$z = w_1x_1 + w_2x_2 + \dots + w_nx_n + b \quad (1)$$

Berdasarkan persamaan, dapat diperoleh output dengan *activation functions*. *ANN* terdiri dari *input layer*, *hidden layers* dan *output layer*. Fungsi aktivasi yang digunakan pada *hidden layer* yaitu *ReLU*. *ReLU* mengubah nilai *output* dalam kisaran 0 (nol) hingga *infinity* dan mengubah nilai negatif menjadi 0.

$$a = \begin{cases} z & \text{if } z \geq 0 \\ 0 & \text{if } z \leq 0 \end{cases} \quad (2)$$

*Softmax* adalah fungsi aktivasi yang digunakan pada *output layer*. *Softmax* memperoleh probabilitas setiap *class* yang jika dijumlahkan harus bernilai 1. Ketika probabilitas suatu *class* meningkat, maka probabilitas *class* lain menurun. Probabilitas tertinggi merupakan *output class*.

$$a_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (3)$$

*Artificial Neural Network* dalam penelitian ini menggunakan metode *optimization Adam*. Dalam persamaan (4), *Adam* menggunakan  $\beta_1$  dan  $\beta_2$  untuk menghitung rata-rata bergerak.  $\beta_1$  untuk perkiraan *moment* pertama dengan nilai *default* 0.9.  $\beta_2$  untuk perkiraan *moment* kedua dengan nilai *default* 0.999.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \quad (4)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) g_t^2 \quad (5)$$

Algoritma yang bergantung pada rata bergerak sebagian besar mengalami bias, sehingga perlu tahap *bias correction step*.

$$\hat{m}_t = \frac{m_t}{1 - \beta_1^t} \quad (6)$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \quad (7)$$

Berdasarkan *biases correction*, dapat dilakukan pembaruan nilai *weights* dan *biases* terhitung dengan nilai *step size*  $\eta$ . Perhitungan parameter *weight* dan bias menggunakan persamaan (8).

$$w_t = w_{t-1} - \eta \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \quad (8)$$

Penelitian ini akan memodifikasi *Adam* dengan memanfaatkan *Nesterov (NAG)*. *Nadam* merupakan algoritma yang menggabungkan *Adam* dan *Nesterov Momentum*. Modifikasi *Adam* hanya dilakukan pada  $\hat{m}_t$  dan tidak pada  $\hat{v}_t$ . Perhitungan  $\hat{m}_t$  menggunakan persamaan (10). Sebelumnya, modifikasi inisialisasi dilakukan untuk memperbarui *gradient*. Persamaan (11) untuk memperbarui nilai *weight* sesuai  $\hat{m}_t$ .

$$g_t = \nabla_{w_t} J(w_t) \quad (9)$$

$$m_t = \gamma m_{t-1} + \eta g_t \quad (10)$$

$$w_{t-1} = w_t - (\gamma m_t + \eta g_t) \quad (11)$$

Evaluasi model menggunakan *multiclass cross entropy* dan *confusion matrix*. Perhitungan nilai *loss* secara terpisah untuk setiap *classes* dilakukan *cross entropy* menggunakan persamaan (12).

$$L = - \sum_{c=1}^M y_c \log(p_c) \quad (12)$$

*Confusion matrix* digunakan untuk menghitung tingkat akurasi model. Akurasi menggambarkan nilai perbandingan jumlah prediksi benar dan semua data masukan menggunakan persamaan (13).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (13)$$

Pelatihan model dilaksanakan dengan beberapa skenario untuk menemukan arsitektur dengan kinerja terbaik. Skenario parameter yang digunakan dalam penelitian ini ditunjukkan pada Tabel 2.

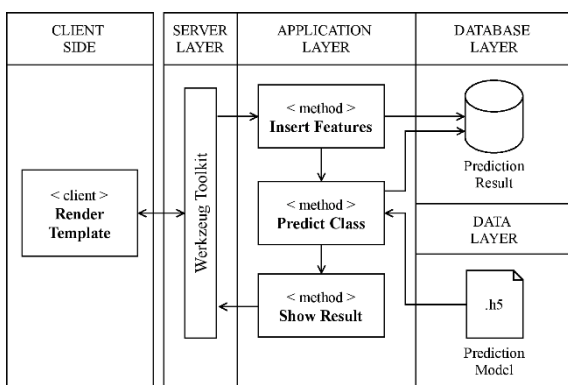
Tabel 2. Parameter Model *Artificial Neural Network*

Parameter	Value
Number of Input	7
Number of Layers	2
Number of Neuron	6,8,10,12,14,16,18,20,22 (trial & error)
Number of Output	1
Hidden Activation	ReLU
Output Activation	Softmax
Optimizer	Adam, Adam + Nesterov (trial & error)
Loss Function	Cross Entropy
Batch Size	32
Epoch	500, 1000 (trial & error)

Berdasarkan parameter model *ANN* yang telah ditentukan, terdapat 36 *training scenarios*. Pengujian model dilakukan menggunakan skenario pembagian data *5-fold cross validation*.

### 2.3. Deployment

*Deployment* meliputi proses penyimpanan dan penerapan model *ANN* untuk memprediksi kesiapan sekolah. Penyimpanan *ANN* dengan *dump function* pada *pickle*. Model yang tersimpan dalam format tertentu, kemudian diterapkan dalam pengembangan sistem web prediksi menggunakan fungsi *load model*. Pengembangan sistem prediksi menggunakan *Flask* dengan arsitektur sistem sebagai berikut.



Gambar 2. Arsitektur Sistem Prediksi

## 3. HASIL DAN PEMBAHASAN

### 3.1. Hasil Eksplorasi Data

Penelitian ini menggunakan *Chi-Square* untuk mengetahui hubungan antara variabel *feature* dengan variabel target yaitu kesiapan sekolah. Hasil korelasi ditunjukkan pada Tabel 3.

Tabel 3. Hubungan Variabel Fitur dengan Variabel Target

Variable		Chi <sup>2</sup>	Sig.
Age	↔ School Readiness	102.32	.000
Gender	↔ School Readiness	2.11	.349
Child Position	↔ School Readiness	20.50	.009
Number of Siblings	↔ School Readiness	128.54	.000
Have preelementary	↔ School Readiness	187.16	.000
Father Education	↔ School Readiness	37.37	.000
Mother Education	↔ School Readiness	23.22	.001

Correlation is significant at the 0.05 level

Berdasarkan hasil korelasi, terdapat hubungan signifikan kuat antara status prasekolah dan kesiapan sekolah dengan nilai  $\chi^2=187.16$  dan  $p<.05$ . Jumlah saudara juga memiliki hubungan *significant* yang kuat dengan nilai  $\chi^2=128.54$  dan  $p<.05$ . Hubungan *significant* yang kuat juga terjadi pada variabel umur dengan nilai  $\chi^2=102.32$ ,  $p<.05$ .

Namun, terdapat hubungan *significant* lemah antara pendidikan ayah, pendidikan ibu, urutan anak dengan kesiapan sekolah dengan  $\chi^2$  20.50 - 37.37 dan  $p<.05$ . Jenis kelamin dan kesiapan sekolah tidak memiliki hubungan signifikan dengan  $\chi^2=2.11$  dan  $p>.05$  (*no significant*).

### 3.2. Hasil Akurasi

Berdasarkan hasil akurasi pada Tabel 4, tingkat akurasi tertinggi adalah sebesar 0.96 didapatkan oleh metode kombinasi *Adam* dan *Nesterov Accelerated Gradient* dengan *architecture* model 10 *neurons* pada *hidden layer*, 1000 *epochs*. Akurasi terendah sebesar 0.8533 diperoleh metode *Adam* dengan arsitektur 6 *neurons* pada *hidden layer*, 500 *epochs*.

Tabel 4. Hasil Akurasi *Adam* dan *Adam + Nesterov*

Architecture	Epoch	Accuracy	
		Adam	Adam + Nesterov
7-6-6-1	500	0.8533	0.8889
7-6-6-1	1000	0.9244	0.9467
7-8-8-1	500	0.8978	0.9111
7-8-8-1	1000	0.9333	0.9289
7-10-10-1	500	0.9111	0.9244
7-10-10-1	1000	0.9511	0.9600
7-12-12-1	500	0.9467	0.9200
7-12-12-1	1000	0.9511	0.9511
7-14-14-1	500	0.9556	0.9467
7-14-14-1	1000	0.9511	0.9511
7-16-16-1	500	0.9289	0.9467
7-16-16-1	1000	0.9422	0.9467
7-18-18-1	500	0.9467	0.9556
7-18-18-1	1000	0.9467	0.9422
7-20-20-1	500	0.9467	0.9467
7-20-20-1	1000	0.9467	0.9511
7-22-22-1	500	0.9422	0.9467
7-22-22-1	1000	0.9556	0.9556

Hasil *accuracy* pada Tabel 4 menunjukkan bahwa perubahan nilai *epoch* memiliki pengaruh terhadap nilai akurasi. Semakin besar nilai *epochs*, maka akan semakin tinggi nilai *accuracy* yang akan didapatkan. Namun, peningkatan jumlah *neurons* dalam *hidden layers* tidak memiliki pengaruh terhadap peningkatan nilai *accuracy*.

### 3.2. Hasil Kecepatan

Evaluasi juga dilakukan menggunakan ukuran waktu atau kecepatan. Waktu yang dimaksud adalah lamanya model melakukan pelatihan dan pengujian. Tabel 5 menunjukkan hasil model *mean time* yaitu jumlah waktu rata-rata untuk seluruh data *fold* pada setiap skenario parameter.

Waktu tercepat dalam melakukan *training* 1,33 *seconds* diperoleh metode *Adam* dengan *architecture* model 6 *neuron* dan 500 *epoch*. Namun, *architecture* tersebut memiliki tingkat akurasi paling rendah yaitu sebesar 0.8533. Sehingga, *architecture* tersebut dapat dikatakan belum optimal.

Tabel 5. Hasil *Modeling Time Adam* dan *Adam + Nesterov*

Architecture	Epoch	Time (second)	
		Adam	Adam + Nesterov
7-6-6-1	500	1.3313	1.3799
7-6-6-1	1000	2.5469	2.6157
7-8-8-1	500	1.3754	1.3638
7-8-8-1	1000	2.3376	2.2251
7-10-10-1	500	1.3656	1.3719
7-10-10-1	1000	2.3876	1.9881
7-12-12-1	500	1.4406	1.4250
7-12-12-1	1000	2.0813	2.1813
7-14-14-1	500	1.4625	1.4719
7-14-14-1	1000	1.9438	1.8844
7-16-16-1	500	1.5219	1.5219
7-16-16-1	1000	1.8844	1.8313
7-18-18-1	500	1.5313	1.4875
7-18-18-1	1000	1.6375	1.6000
7-20-20-1	500	1.6356	1.7449
7-20-20-1	1000	1.7657	1.5938
7-22-22-1	500	1.5250	1.5375
7-22-22-1	1000	1.6313	1.6063

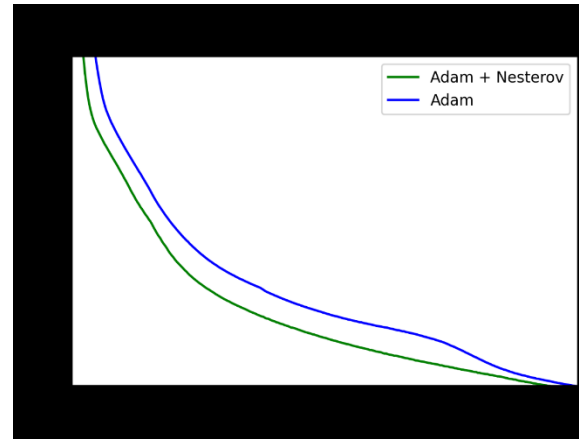
Dalam evaluasi sebelumnya, *accuracy* tertinggi diperoleh kombinasi *Adam* dan *Nesterov's* dengan *architecture* model 10 *neurons* pada *hidden layer* dan 1000 *epochs*. Dalam evaluasi waktu, model tersebut juga memiliki waktu yang relatif cepat dalam proses pemodelan yaitu 1,9881 *seconds*. Sehingga, model tersebut dapat dikatakan optimal.

### 3.3. Hasil Loss

*Loss function* digunakan untuk mengetahui nilai *error* dengan menghitung perbedaan antara *predicted output* dan *desired output*. *Loss function* mengalami peningkatan apabila probabilitas prediksi semakin menyimpang dari *actual output*. Semakin rendah *loss function*, semakin baik model klasifikasi. Hasil *loss function* terbaik adalah bernilai 0.

Evaluasi kombinasi *Adam* dengan *Nesterov's* menggunakan skenario 10 *neuron* pada *hidden layer*. Proses evaluasi dilakukan hingga 1000 *epoch*, namun sebelum mencapai iterasi 1000, model telah berhasil mencapai minimum *loss* pada iterasi ke 300.

Hasil menunjukkan bahwa *loss function* paling rendah didapatkan oleh algoritma kombinasi *Adam* dan *Nesterov Accelerated Gradient*. *Optimization* ini memiliki *loss function* lebih rendah daripada *Adam* pada setiap *epochs*. Minimum *loss* juga lebih cepat didapatkan oleh kombinasi *Adam* dan *Nesterov*.



Gambar 3. *Loss Function Adam* dan *Adam + Nesterov*

Tabel 6. Hasil *Loss Function Adam* dan *Adam + Nesterov*

Optimizer	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean
Adam	0.05	0.06	0.09	0.08	0.06	0.068
Adam+ Nesterov	0.08	0.04	0.06	0.06	0.07	0.062

Berdasarkan hasil *loss* pada Tabel 5, nilai *loss* terendah sebesar 0.04 didapatkan oleh metode *Adam* dengan *Nesterov's Accelerated Gradient*. Metode optimasi ini memiliki *mean loss function* lebih baik daripada *Adam* yaitu senilai 0.062 untuk kombinasi *Adam* dengan *Nesterov* dan 0.068 untuk *Adam*.

### 3.4. Hasil Arsitektur Model

Hasil evaluasi *classical adam* dan *Adam* dengan *Nesterov Momentum* menggunakan nilai akurasi dan *loss function* menunjukkan kinerja terbaik dilakukan oleh kombinasi *Adam* dengan *Nesterovs Momentum*.

Metode optimasi kombinasi *Adam* dan *Nesterov Momentum* menghasilkan nilai *accuracy* 96.00% dan *loss* 0.062 dalam 1.98 *seconds* dengan *architecture* 10 *neuron* pada *hidden layer* dan 1000 *epochs*.

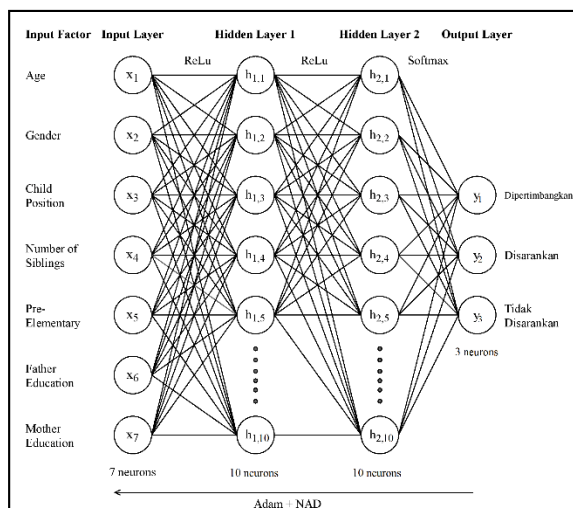
Sedangkan, *classical Adam* atau *Adam* dengan *vanila Momentum* menghasilkan nilai akurasi sebesar 95.00% dan *loss* 0.068 dalam 2.38 *seconds*. Sehingga, dapat diketahui bahwa metode kombinasi *Adam* dan *Nesterovs* memiliki kinerja lebih baik dibandingkan dengan *classical Adam*. Arsitektur model *ANN* dengan kinerja terbaik dapat ditunjukkan pada Gambar 4.

### 3.1. Hasil Deployment

*Deployment* adalah proses penyaluran *machine learning* untuk dimanfaatkan dalam kehidupan. *Flask framework* digunakan dalam pengembangan sistem prediksi kesiapan sekolah berbasis *web*.

Gambar 5 adalah tampilan hasil prediksi. Hasil prediksi menampilkan klasifikasi berupa disarankan, dipertimbangkan atau tidak disarankan untuk masuk sekolah dasar berdasarkan *input* yang telah diterima.





Gambar 4. Arsitektur Model Terbaik

Prediksi Kesiapan Sekolah	
Nama	: SAMPLE 001
Kelas	: TK-B
Jenis Kelamin	: Perempuan
Umur	: 6 Tahun
Anak Ke	: 1
Saudara	: 1
Pendidikan Ayah	: Sarjana
Pendidikan Ibu	: Sarjana
<div style="border: 1px solid green; padding: 5px; margin-top: 10px;"> <b>HASIL PREDIKSI</b>  Disarankan </div>	

Gambar 5. Tampilan Hasil Prediksi Kesiapan Sekolah

Sistem ini dapat menangkap masukan *features*, memuat file model prediksi yang telah tersimpan, melakukan prediksi, menyimpan hasil prediksi pada *database* dan menampilkan riwayat prediksi kepada pengguna. Aktivitas pengguna terdiri dari melakukan prediksi dan menampilkan *history*.

#### 4. KESIMPULAN

Kombinasi *Adam* dengan *Nesterov Momentum* memiliki kinerja lebih baik dibandingkan *classical Adam* dalam melakukan prediksi kesiapan sekolah. Kombinasi metode tersebut menghasilkan *accuracy* sebesar 96.00% dan *loss function* senilai 0.06 dalam 1.98 seconds dengan 10 neurons pada 2 hidden layer dan 1000 epochs.

Selanjutnya, prediksi kesiapan sekolah dapat ditingkatkan dengan menambahkan *features* lainnya, misalnya *family size*, *occupation*, *neighborhood* dan *class peers*. Metode *optimization* dapat ditingkatkan dengan *Accelerate Adam* atau *AAdam* yang mampu meningkatkan kinerja *Adam* serta kombinasi *Adam* dan *Nesterov Momentum*.

#### DAFTAR PUSTAKA

- CHERIDITO, P., JENTZEN, & ROSSMANNEK, F. 2021. Non-convergence of Stochastic Gradient Descent in The Training of Deep Neural Network. *Journal of Complexity*, 64(1), pp.1–10.
- DOZAT, T. 2016. Incorporating Nesterov Momentum Into Adam. *ICLR*.
- FITRI, R., & REZA, M. 2020. Instrumen Kesiapan Belajar: Asesmen Non-Tes untuk Mengukur Kesiapan Belajar Anak Usia Dini dalam Perspektif Neurosains. *Jurnal Pendidikan, Pengasuhan, Kesehatan Dan Fizi Anak Usia Dini*, 1(1), pp.17–32.
- GUNA, P. W. T., & PUTRI, L. A. A. R. 2020. Endek Classification Based on GLCM Using Artificial Neural Network with Adam Optimization. *Jurnal Elektronik Ilmu Komputer Udayana*, 9(2).
- HERDYANA, E. 2019. Perbedaan Masa Perkembangan Anak Prasekolah Usia 48-60 Bulan Berdasarkan Jenis Kelamin dengan Menggunakan Instrumen Kuesioner Pra Skrining Perkembangan (KPSP). *Jurnal Kebidanan*, 8(1), pp.40–46.
- HUDA, N. S., MUBAROK, M. S., & ADIWIJAYA. 2019. A Multi-label Classification on Topics of Quranic Verses (English Translation) Using Backpropagation Neural Network with Stochastic Gradient Descent and Adam Optimizer. In *International Conference on Information and Communication Technology*. Kuala Lumpur: IEEE.
- INTAN, GHANI, & SALMAN. 2019. Implementation of The K-Nearest Neighbor and Neural Network for Predicting School Readiness to Enter Elementary School. In *International Conference on Cyber and IT Service Management (CITSM)*. Jakarta: IEEE.
- KINGMAN, D. P., & BA, J. 2015. Adam: A Method for Stochastic Optimization. In *3rd International Conference for Learning Representations*. San Diego.
- MENDIKBUD. 2021. Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 1 Tahun 2021 Tentang Penerimaan Peserta Didik Baru pada Taman Kanak-Kanak, Sekolah Dasar, Sekolah Menengah Pertama, Sekolah Menengah Atas dan Sekolah

- Menengah Kejuruan. Jakarta: Kementerian Pendidikan dan Kebudayaan Republik Indonesia.
- NURLINA, INTAN, I., & SYAHRIANI. 2019. Pre Test Prediction System for Preparing Readiness for Basic Education. In Seminar Nasional Komunikasi dan Informatika, pp.119–124. Jakarta: Jurnal Kominfo.
- SHAW, D., MENDELSON, A., & MORRIS, P. 2021. Reducing Poverty-Related Disparities in Child Development and School Readiness: The Smart Beginnings Tiered Prevention Strategy that Combines Pediatric Primary Care with Home Visiting. *Clinical Child and Family Psychology Review*, (24), pp.669–683.
- SUSILARINI, T. 2021. Deteksi Dini Kesiapan Memasuki Sekolah Dasar Melalui Nijmeegse Schoolbekwaamheid Test (NST) dan Children Personal Matrix Test (CPM). *Jurnal IKRA-ITH Humaniora*, 5(1), pp.177–182.
- WANG, Y., ZHOU, P., & ZHONG, W. 2018. An Optimization Strategy Based on Hybrid Algorithm of Adam and SGD. In 2nd International Conference on Electronic Information Technology and Computer Engineering. Shanghai.
- WARDHANI, R. K., JAYA, S. T., & FAUZIYAH, N. 2019. Hubungan Pengetahuan Ibu Tentang Stimulasi Perkembangan dengan Perkembangan Personal Sosial pada Anak. *Jurnal Ilmiah Pamenang*, 1(1), pp.39–44.
- WULANSUCI, G., & KURNIATI, E. 2019. Pembelajaran Calistung (Membaca, Menulis, Berhitung) dengan Resiko Terjadinya Stress Akademik pada Anak Usia Dini. *Jurnal Tunas Siliwangi*, 5(1).

*Halaman ini sengaja dikosongkan*