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

SKRIPSI

UIN SUNAN AMPEL
S U R A B A Y A

DISUSUN OLEH:

INDAH MUSTIKA RAHAYU

H76217059

PROGRAM STUDI SISTEM INFORMASI

FAKULTAS SAINS DAN TEKNOLOGI

UNIVERSITAS ISLAM NEGERI SUNAN AMPEL

SURABAYA

2021

LEMBAR PERNYATAAN KEASLIAN

Saya yang bertanda tangan dibawah ini,

NAMA	:	INDAH MUSTIKA RAHAYU
NIM	:	H76217059
PROGRAM STUDI	:	SISTEM INFORMASI
ANGKATAN	:	2017

Menyatakan bahwa saya tidak melakukan plagiasi dalam penulisan skripsi saya yang berjudul: "PREDIKSI KESIAPAN SEKOLAH MENGGUNAKAN *MACHINE LEARNING* BERBASIS KOMBINASI *ADAM* DAN *NESTEROV MOMENTUM*". Apabila suatu saat nanti terbukti saya melakukan tindakan plagiat, maka saya bersedia menerima sanksi yang telah ditetapkan.

Demikian pernyataan keaslian ini saya buat dengan sebenar-benarnya.

Surabaya, 12 Agustus 2021

Yang Menyatakan,

Indah Mustika Rahayu

NIM. H76217059

LEMBAR PERSETUJUAN PEMBIMBING

Skripsi oleh

NAMA : INDAH MUSTIKA RAHAYU

NIM : H76217059

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

Ini telah diperiksa dan disetujui untuk diujikan

Surabaya, 5 Agustus 2021

Menyetujui,

Dosen Pembimbing 1	Dosen Pembimbing 2
Mujib Ridwan, M.T	Ahmad Yusuf, M.Kom
NIP. 198604272014031004	NIP. 199001202014031003

PENGESAHAN TIM PENGUJI SKRIPSI

Skripsi Indah Mustika Rahayu ini telah dipertahankan

di depan tim penguji skripsi

di Surabaya, 10 Agustus 2021

Mengesahkan,

Dewan Penguji

Penguji I	Penguji II
Mujib Ridwan, M.T	Ahmad Yusuf, M.Kom
NIP. 198604272014031004	NIP. 199001202014031003

Penguji III	Penguji IV
Khalid, M.Kom	Achmad Teguh Wibowo, M.T
NIP. 197906092014031002	NIP. 198001152014031003

Mengetahui,

Dekan Fakultas Sains dan Teknologi

UIN Sunan Ampel Surabaya

Dr. Evi Fatimatur Rusydiyah, M.Ag

NIP. 197312272005012003

KEMENTERIAN AGAMA
UNIVERSITAS ISLAM NEGERI SUNAN AMPEL SURABAYA
PERPUSTAKAAN
Jl. Jend. A. Yani 117 Surabaya 60237 Telp. 031-8431972 Fax.031-8413300
E-Mail: perpus@uinsby.ac.id

LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS

Sebagai sivitas akademika UIN Sunan Ampel Surabaya, yang bertanda tangan di bawah ini, saya:

Nama	: INDAH MUSTIKA RAHAYU
NIM	: H76217059
Fakultas/Jurusan	: SAINS & TEKNOLOGI / SISTEM INFORMASI
E-mail address	: indahrahayumustika@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan UIN Sunan Ampel Surabaya, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah :
■ Sekripsi ☐ Tesis ☐ Desertasi ☐ Lain-lain (.....)
yang berjudul :
PREDIKSI KESIAPAN SEKOLAH MENGGUNAKAN *MACHINE LEARNING*
BERBASIS KOMBINASI *ADAM* DAN *NESTEROV MOMENTUM*

beserta perangkat yang diperlukan (bila ada). Dengan Hak Bebas Royalti Non-Ekslusif ini Perpustakaan UIN Sunan Ampel Surabaya berhak menyimpan, mengalih-media/format-kan, mengelolanya dalam bentuk pangkalan data (database), mendistribusikannya, dan menampilkan/mempublikasikannya di Internet atau media lain secara ***fulltext*** untuk kepentingan akademis tanpa perlu meminta ijin dari saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan atau penerbit yang bersangkutan.

Saya bersedia untuk menanggung secara pribadi, tanpa melibatkan pihak Perpustakaan UIN Sunan Ampel Surabaya, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini yang saya buat dengan sebenarnya.

Surabaya, 12 Agustus 2021

Penulis

(Indah Mustika Rahayu)
nama terang dan tanda tangan

ABSTRAK

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

Oleh:

Indah Mustika Rahayu

Kesiapan sekolah perlu diukur untuk mengetahui tingkat kematangan anak dan mengetahui kemampuan anak yang perlu dikembangan. Namun, instrumen tes *NST* untuk mengukur kesiapan sekolah hanya dapat digunakan oleh Biro Psikologi, sehingga guru dan orang tua yang berperan dalam bentuk pemberian dukungan dan stimulasi pada anak tidak dapat menggunakan instrumen tersebut. Oleh karena itu, diperlukan *machine learning* untuk menciptakan model prediksi kesiapan sekolah. Data penelitian adalah data administratif anak yang telah mengikuti tes kesiapan sekolah berbasis *NST* yang diselenggarakan oleh TK Ar-Rasyid pada tahun 2012-2018. Penelitian menggunakan algoritma *Artificial Neural Network* dengan metode optimasi kombinasi *Adam* dan *Nesterov Momentum*. Pengujian model mengunakan skenario *5-fold cross validation*. Hasil penelitian menunjukkan bahwa kombinasi *Adam* dan *Nesterov Momentum* memiliki kinerja lebih baik daripada *clasical 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**

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

ABSTRACT

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

***By*:**

Indah Mustika Rahayu

School readiness needs to be measured to know the level of maturity of the child and to know the ability of the child that needs to be developed. However, the NST test instrument to measure school readiness can only be used by the Bureau of Psychology, so teachers and parents who play a role in providing support and stimulation to children cannot use the instrument. Therefore, machine learning is needed to create a prediction model of school readiness. Research data is administrative data of children who have taken the NST-based school readiness test conducted by TK Ar-Rasyid on 2012-2018. The research used Aritificial Neural Network algorithms with 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

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR ISI

LEMBAR PERNYATAAN KEASLIAN ..... **ii**
LEMBAR PERSETUJUAN PEMBIMBING ..... **iii**
PENGESAHAN TIM PENGUJI SKRIPSI ..... **iv**
LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI ..... **v**
DAFTAR ISI ..... **vi**
DAFTAR TABEL ..... **viii**
DAFTAR GAMBAR ..... **ix**
ABSTRAK ..... **x**
ABSTRACT ..... **xi**
BAB I PENDAHULUAN **1**
1.1 Latar Belakang 1
1.2 Perumusan Masalah 4
1.3 Batasan Masalah 4
1.4 Tujuan Penelitian 4
1.5 Manfaat Penelitian 5
1.6 Sistematika Penulisan 5
BAB II TINJAUAN PUSTAKA **6**
2.1 Tinjauan Penelitian Terdahulu 6
2.2 Teori Dasar 7
2.2.1 *Nijmeegse Schoolbekwaamheids Test* 7
2.2.2 *Data Mining* 8
2.2.3 *Machine Learning* 8
2.2.4 Klasifikasi 9
2.2.5 *Artificial Neural Network* 9
2.2.6 *Adam Optimizer* 12
2.2.7 Kombinasi *Adam* dan *Nesterov* 13
2.2.8 *Cross Entropy* 14
2.2.9 *Confusion Matrix* 15

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.2.10 *Python Programming Language* 16
2.2.11 *Scikit Learn Library* 16
2.2.12 *Flask Framework* 16
2.3 Integrasi Keilmuan 17
BAB III METODOLOGI PENELITIAN 18
3.1 Metode Penelitian 18
3.1.1 Perumusan Masalah 18
3.1.2 Studi Pustaka 19
3.1.3 *Data Understanding* 19
3.1.4 *Data Preparation* 21
3.1.5 *Modeling* 21
3.1.6 *Deployment* 24
3.1.7 Evaluasi Hasil 24
BAB IV HASIL DAN PEMBAHASAN 25
4.1 *Data Understanding* 25
4.2 *Data Preparation* 27
4.3 *Modeling* 28
4.3.1 *Model Training* 29
4.3.2 *Model Evaluation* 36
4.4 *Deployement* 41
4.5 Diskusi 47
BAB V PENUTUP 48
5.1 Kesimpulan 48
5.2 Saran 48
DAFTAR PUSTAKA 49
LAMPIRAN 53

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR TABEL

Tabel 2.1 Tinjauan Penelitian Terdahulu 6
Tabel 2.2 Ilustrasi Confusion Matrix 15

Tabel 3.1 Atribut Penelitian Terdahulu..... 19
Tabel 3.2 Deskripsi Fitur dan Label..... 20
Tabel 3.3 Parameter Model Artificial Neural Network 21
Tabel 3.4 Parameter Model Artificial Neural Network lanjutan..... 22
Tabel 3.5 Skenario Pelatihan Model Artificial Neural Network..... 23

Tabel 4.1 Descriptive Statistics Categorical Features..... 25
Tabel 4.2 Descriptive Statistics Numerical Features 26
Tabel 4.3 Jumlah Data Setiap Target Class 26
Tabel 4.4 Hubungan Variable Fitur dengan Variable Target..... 27
Tabel 4.5 Data Formating pada Data Categorical..... 28
Tabel 4.6 Hasil Accuracy ANN dengan Adam dan Adam Nesterov..... 36
Tabel 4.7 Hasil Accuracy ANN dengan Adam dan Adam Nesterov (lanjutan) ... 37
Tabel 4.8 Hasil Modeling Time Adam dan Adam Nesterov 38
Tabel 4.9 Hasil Loss Function ANN dengan Adam dan Adam Nesterov 39
Tabel 4.10 Hasil Pengujian Sistem menggunakan Adam 45
Tabel 4.11 Hasil Pengujian Sistem menggunakan Adam dan Nesterov..... 45
Tabel 4.12 Incorrect Prediction Metode Adam..... 46
Tabel 4.13 Incorrect Prediction Metode Adam Nesterov 46

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR GAMBAR

Gambar 2.1 Single Perceptron Illustration 10

Gambar 3.1 Metode Penelitian 18
Gambar 3.2 Arsitektur Model Artificial Neural Network 22
Gambar 3.3 Diagram Block Model Artificial Neural Network 24

Gambar 4. 1 Hasil Perhitungan Neuron Pada Input Layer 31
Gambar 4.2 Hasil Perhitungan Neuron Pada Hidden Layer 32
Gambar 4.3 Hasil Perhitungan Neuron Pada Ouput Layer 33
Gambar 4.4 Hasil Pembaruan Weight Menggunakan Adam 34
Gambar 4.5 Hasil Pembaruan Weight Menggunakan Adam dan Nesterov 35
Gambar 4.7 Kurva Loss Function Metode Adam dan Adam Nesterov 39
Gambar 4.8 Arsitektur Model Artificial Neural Network Terbaik 40
Gambar 4.9 Arsitektur Sistem Prediksi 41
Gambar 4.10 Activity Diagram Prediksi 42
Gambar 4.11 Activity Diagram History 42
Gambar 4.12 User Interface Input Data 43
Gambar 4.13 User Interface Hasil Prediksi 43
Gambar 4.14 User Interface History Data 44

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB I
PENDAHULUAN

1.1 Latar Belakang

Kesiapan sekolah adalah konsep multidimensi meliputi karakter kognitif, fungsi eksekusi, bahasa, sosial emosional, perilaku serta kesehatan lingkungan yang berkontribusi pada kemampuan anak untuk beradaptasi dan berkembang dalam sistematika sekolah (Sabol & Pianta, 2017). Berdasarkan Peraturan Mendibud, usia yang tepat untuk masuk sekolah dasar adalah tujuh tahun atau paling rendah enam tahun bagi anak yang telah memiliki kesiapan psikis (Mendikbud, 2018).

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

Namun, setiap anak tidak memiliki kondisi fisik dan psikis yang sama pada usia tertentu. Pada usia prasekolah yaitu 3-6 tahun, anak perempuan lebih dapat menyesuaikan diri dengan lingkungan sosial dibandingkan dengan anak laki-laki (Wardhani, Jaya, & Fauziyah, 2019). Jenis kelamin dapat memengaruhi perbedaan masa perkembangan anak (Herdyana, 2019). Sehingga, untuk memastikan kesiapan anak untuk masuk pendidikan dasar, diperlukan *Nijmeegse Schoolbekwaamheids Test* (*NST*) (Monks, Rost, & Coffie, 1978).

Nijmeegse Schoolbekwaamheids Test adalah instrumen untuk mengukur kesiapan sekolah yang meliputi kesiapan fisik dan psikis. Tes ini dilakukan dengan menyajikan beberapa gambar untuk dideskripsikan atau dilengkapi. Hasil tes akan menunjukkan karakteristik kognitif, motorik, sosial dan emosional yang dimiliki anak. Berdasarkan hasil tersebut, dapat diketahui bahwa anak siap atau tidak untuk memasuki jenjang sekolah dasar (Fauziyah, 2019).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Kesiapan sekolah perlu diukur untuk mengetahui tingkat kematangan anak dan mengetahui kemampuan anak yang perlu dikembangan. Namun, instrumen tes *NST* untuk mengukur kesiapan sekolah hanya dapat digunakan oleh Biro Psikologi, sehingga guru dan orang tua yang berperan dalam bentuk pemberian dukungan dan stimulasi pada anak tidak dapat menggunakan instrumen tersebut (Fitri & Reza, 2020). Oleh karena itu, diperlukan *data mining* yang akan memudahkan guru dan orang tua untuk memprediksi kesiapan sekolah (Intan, Ghani, & Salman, 2019).

Data mining adalah proses menganalisi dan mengektrak pengetahuan secara iteratif menggunakan *machine learning* untuk menemukan pola yang berguna dalam data yang besar (Gorunescu, 2011). *Machine learning* didefinisikan sebagai penggunaan algoritma dan statistik untuk belajar dari data tanpa diprogram secara eksplisit. *Data mining* telah berkontribusi dalam studi pendidikan tentang cara siswa belajar untuk meningkatkan praktik dan materi pembelajaran (Calders, 2012).

Data anak dalam proses administrasi *Nijmeegse Schoolbekwaamheids Test* yang terdiri dari umur, jenis kelamin, urutan anak antara saudara kandung, jumlah saudara, status pendidikan prasekolah, pendidikan ayah dan pendidikan ibu dapat dimanfaatkan untuk melakukan prediksi kesiapan sekolah (Nurlina et al., 2019). Algoritma klasifikasi memiliki kinerja yang baik dalam memprediksi kesiapan anak untuk sekolah (Suleiman, Arslan, & Ridley, 2017).

Algoritma klasifikasi yang telah digunakan dalam memprediksi kesiapan sekolah yaitu *Decision Tree*, *K-Nearest Neighbor* dan *Artificial Neural Network* (*ANN*) (Suleiman et al., 2017) (Nurlina et al., 2019) (Intan et al., 2019). Algoritma *Artificial Neural Network* dengan metode *Stochastic Gradient Descent* mempunyai akurasi lebih baik dibandingkan algoritma *Decision Tree* dan *K-Nearest Neighbor* dalam melakukan klasifikasi kesiapan sekolah (Intan et al., 2019).

Artificial Neural Network juga memiliki kinerja yang lebih baik daripada *Support Vector Machine*, *Naïve Bayes*, *K-Nearest Neighbor*, *Decision Tree* dan *Random Forest* dalam memprediksi penyakit diabetes dengan nilai akurasi 88,71% (Pradhan, Dhaka, & Kulhari, 2020). Selain itu, dalam deteksi kantuk pengemudi, algoritma *ANN* juga mempunyai akurasi lebih baik dibandingkan *K-NN* dengan akurasi *ANN* sebesar 79.33% dan *K-NN* sebesar 72,19% (Li, Zhang, & Zhao, 2017).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Artficial Neural Network adalah teknik pemodelan data numeric non-linear yang dapat memodelkan hubungan yang kompleks di antara input dan output. *ANN* juga memiliki kemampuan generalisasi yang dapat mengetahui hubungan yang tidak terlihat pada data baru (El-Khatib, 2019). *Artificial Neural Network* memiliki kinerja yang baik dalam memodelkan data non-linear dengan jumlah input yang besar, sehingga model ini sesuai dengan model prediksi kesiapan sekolah.

Metode optimasi model *ANN* dengan *Stochastic Gradient Descent* (*SGD*) telah memiliki kinerja yang baik dalam berbagai aplikasi, namun *SGD* menetapkan skala gradien secara seragam ke segala arah. Hal tersebut dapat merugikan kinerja *ANN* jika skala gradien kecil (Wang, Zhou, & Zhong, 2018). Untuk memperbaiki kekurangan *SGD*, metode adaptif telah diusulkan yang secara diagonal menetapkan skala dari gradien melalui perkiraan kelengkungan fungsi. Metode tersebut adalah *Adaptive Moment Estimation* atau *Adam* (Kingman & Ba, 2015).

Adam Optimizer memiliki kinerja lebih baik daripada *SGD* (Guna & Putri, 2020). Metode optimasi *Adam* juga memiliki kemampuan yang lebih cepat dalam proses *training* dibandingkan dengan metode *SGD* (Huda, Mubarok, & Adiwijaya, 2019). Selain itu, dalam studi terkait perbandingan metode optimasi *ANN* dalam melakukan klasifikasi pada berbagai *datasets* menunjukkan bahwa *Adam* memiliki akurasi lebih baik daripada metode *SGD*, *Nesterov Momentum*, *Adagrad*, *Adadelta* dan *RMSProp* dengan rata-rata nilai akurasi 98% (Zaheer & Shaziya, 2019).

Artificial Neural Network dengan *Adam Optimizer* dapat digunakan sebagai solusi untuk meningkatkan kinerja model prediksi kesiapan sekolah. Metode *Adam* memiliki komponen momentum dan komponen *adaptive learning rate*. Komponen momentum yang digunakan oleh *Adam* adalah *vanilla momentum*, sehingga pada segi momentum metode *Nesterov Momentum* lebih baik daripada *Adam*.

Adam optimizer 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 nilai *loss* karena osilasi yang besar. Untuk menghindari osilasi, gradien akan diarahkan pada parameter sebelumnya. Hal tersebut yang dilakukan oleh *Nesterov Momentum*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Artificial Neural Network dengan metode kombinasi *Adam Optimizer* serta *Nesterov Momentum* adalah algoritma *machine learning* yang diharapkan mampu meningkatkan kinerja model prediksi kesiapan sekolah. Berdasarkan latar belakang yang telah diuraikan, maka perlu dilaksanakan penelitian dengan judul "**PREDIKSI KESIAPAN SEKOLAH MENGGUNAKAN *MACHINE LEARNING* BERBASIS KOMBINASI *ADAM* DAN *NESTEROV MOMENTUM***", sehingga penelitian dapat dimanfaatkan oleh guru dan orang tua untuk mengetahui kesiapan sekolah anak.

1.2 Perumusan Masalah

1. Bagaimana kinerja *machine learning* berbasis kombinasi *Adam* daan *Nesterov Momentum* dibandingkan dengan *machine learning* berbasis *Adam* dalam memprediksi kesiapan sekolah?
2. Bagaimana ketepatan metode kombinasi *Adam* dan *Nesterov Momentu* dibandingkan dengan *Adam* dalam memprediksi kesiapaan sekolah sesuai dengan hasil *NST*?

1.3 Batasan Masalah

Batasan ruang lingkup dalam penelitian ini adalah sebagai berikut.

1. Data penelitian adalah data administrasi beserta hasil tes dari anak yang telah mengikuti *Nijmeegse Schoolbekwaamheids Test*.
2. Data penelitian didapatkan dari TK Ar-Rasyid sebagai penyelenggara.
3. Penelitian menggunakan data pada tahun 2012-2018.
4. Arsitektur *ANN* menggunakan 2 *hidden layers*.
5. Penelitian menggunakan batuan *scikit-learn* dan *flask*.

1.4 Tujuan Penelitian

Berdasarkan rumusan masalah, didapatkan tujuan penelitian berikut.

1. Mengetahui kinerja *machine learning* berbasis kombinasi *Adam* daan *Nesterov Momentum* dibandingkan dengan *machine learning* berbasis *Adam* dalam memprediksi kesiapan sekolah.
2. Mengetahui ketepatan metode kombinasi *Adam* dan *Nesterov Momentu* dibandingkan dengan *Adam* dalam memprediksi kesiapan sekolah sesuai dengan hasil *NST*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

1.5 Manfaat Penelitian

1. Manfaat Teoritis
   a. Mengembangkan model prediksi kesiapan sekolah menggunakan *machine learning* dengan algoritma *Artificial Neural Network*.
   b. Meningkatkan kinerja *machine learning* dengan metode kombinasi *Adam* dan *Nesterov Momentum*.
2. Manfaat Praktis
   a. Memudahkan guru untuk mengetahui kesiapan anak untuk sekolah, sehingga dapat dijadikan sebagai bahan evaluasi.
   b. Memudahkan orang tua untuk mengetahui kesiapan sekolah anak, sehingga dapat dimanfaatkan sebagai bahan pertimbangan untuk mengikutsertakan anak pada jenjang pendidikan Sekolah Dasar.

1.6 Sistematika Penulisan

Sistematika penulisan dalam penelitian dapat dipaparkan sebagai berikut.

1. BAB I PENDAHULUAN

   Pendahuluan memaparkan latar belakang penelitian, rumusan masalah, batasan masalah, tujuan penelitian dan manfaat penelitian.
2. BAB II TINJAUAN PUSTAKA

   Tinjauan pustaka akan menjelaskan tinjauan penelitian terdahulu, teori dasar terkait *Adam*, *Nesterov Momentum* serta integrasi keilmuan.
3. BAB III METODOLOGI PENELITIAN

   Metodologi penelitian memaparkan tahapan dalam mewujudkan hasil penelitian. Hal tersebut juga terkait dengan variable dan *tools* yang akan dimanfaatkan untuk melakukan penelitian.
4. BAB IV HASIL DAN PEMBAHASAN

   Hasil dan pembahasan akan menjelaskan kinerja kombinasi *Adam* dan *Nesterov Momentum* dalam memprediksi kesiapan sekolah.
5. BAB V PENUTUP

   Penutup menjelaskan kesimpulan dengan menguraikan hasil penelitian berdasarkan rumusan masalah serta menjelaskan saran pengembangan berdasarkan hasil penelitian yang telah didapatkan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB II
TINJAUAN PUSTAKA

2.1 Tinjauan Penelitian Terdahulu

Tinjauan penelitian terdahulu dilakukan dengan merangkum penelitian yang memiliki relasi. Hasil penelitian terdahulu mampu berfungsi sebagai data referensi. Selain itu, dapat mengetahui penelitian yang perlu adanya pengembangan. Tabel 2.1 menunjukkan tinjauan penelitian terdahulu.

Tabel 2.1 Tinjauan Penelitian Terdahulu

No	Judul	Hasil	Korelasi
1	*Prediction Model for School Readiness* (Suleiman et al., 2017)	Algoritma klasifikasi *J48* dan *ID3* dapat dengan baik melakukan prediksi kesiapan anak untuk sekolah.	Menggunakan *data mining* untuk membantu memprediksi kesiapan sekolah.
2	*Pre Test Prediction System for Preparing Readiness for Basic Education* (Nurlina et al., 2019)	Sistem prediksi kesiapan sekolah dengan menggunakan algoritma *K-Nearest Neighbor* menghasilkan akurasi sebesar 68.3% dengan nilai k = 5.	
3	*Implementation of The K-Nearest Neighbor for Predicting School Readiness to Enter Elementary School* (Intan et al., 2019)	Algoritma *Neural Network* menghasilkan akurasi lebih baik daripada *K-Nearest Neighbor*. *KNN* menghasilkan akurasi sebesar 75.84% dengan nilai k = 9. *NN* menghasilkan akurasi sebesar 82.01% dengan *neuron* = 8, *learning rate* = 0.8, *training cycle* = 500 dan *momentum* = 0.8.	Menggunakan *artificial neural network* untuk membantu memprediksi kesiapan sekolah.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4	*Incorporating Nesterov Momentum Into Adam* (Dozat, 2016)	Modifikasi komponen *Adam* dengan memanfaatkan *Nesterov Momentum* dapat meningkatkan kinerja model.	Menggunakan *adam optimizer* untuk membantu meningkatkan kinerja *artificial neural network*.
5	*Classification on Topics of Quranic Verses Using Neural Network with SGD and Adam Optimizer* (Huda et al., 2019)	*Adam optimizer* menghasilkan *training accuracy* lebih besar daripada *SGD* di awal *epoch*. Namun, *training* yang lebih baik tidak selalu memberikan *validation* yang lebih baik.	
6	*Endek Classification Based Using Artificial Neural Network with Adam Optimizer* (Guna & Putri, 2020)	*Adam algorithm* menghasilkan akurasi lebih baik daripada *SGD Algorithm*. Pada nilai *epoch* yang rendah, *adam algorithm* juga menghasilkan akurasi yang baik.	

Berdasarkan tinjauan penelitian terdahulu, dapat diketahui bahwa *Artificial Neural Network* memiliki akurasi lebih tinggi daripada algoritma *Decision Tree* dan *K-Nearest Neighbor* dalam melakukan prediksi kesiapan sekolah. Selain itu, dapat diketahui bahwa model *Artificial Neural Network* dapat ditingkatkan dengan *Adam Optimizer*. Sehingga, penelitian ini dilakukan untuk menerapkan *Adam Optimizer* dalam model *ANN* untuk meningkatkan kinerja prediksi kesiapan sekolah.

2.2 Teori Dasar

2.2.1 *Nijmeegse Schoolbekwaamheids Test*

Nijmeegse Schoolbekwaamheids Test adalah instrumen tes untuk mengukur kesiapan sekolah. Pelaksanaan tes dilakukan dengan menyajikan beberapa gambar untuk dideskripsikan atau dilengkapi. Instrumen ini terdiri dari sepuluh aspek yang akan diukur untuk mengetahui kesiapan sekolah. Aspek tersebut yaitu pengamatan bentuk, motorik halus, perbandingan, pengaamatan kritis, konsentrasi, daya ingat, penilaian situasi, memahami cerita serta gambar orang (Monks et al., 1978).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Aspek memiliki beberapa materi soal dan memiliki bobot untuk setiap soal. Hasil setiap aspek kemudian akan dipetakan pada kolom profil yang masuk dalam norma kesiapan. Hasil akan menunjukkan karakteristik kognitif, motorik, sosial dan emosional yang dimiliki anak. Berdasarkan hasil tersebut, dapat diketahui bahwa anak dapat dikategorikan dalam kelompok disarankan, dipertimbangkan atau tidak disarankan untuk memasuki jenjang pendidikan sekolah dasar (Fauziyah, 2019).

2.2.2 *Data Mining*

Data mining yaitu prosedur ekstraksi pola dan satu model yang berguna dari himpunan data yang besar. Model dan pola memiliki peran yang efektif dalam tugas pengambilan keputusan. *Data mining* adalah langkah dari *Knowledge Discovey in Database* yang melakukan analisis dan pemodelan pada kumpulan data yang besar menggunakan *machine learning* terdiri dari metode untuk *clasification*, *estimation*, *forecasting*, *clustering* atau *association rule* (Alasadi & Bhaya, 2017).

Knowledge Discovey Step meliputi *data selection* untuk menentukan data target, *data preprocessing* untuk meningkatkan keandalannya, *data transformation* untuk mengubah data dalam format tertentu untuk proses penambangan data, *data mining* dengan algoritma *classification*, *estimation*, *forecasting*, *clustering*, atau *association rule* untuk mengekstrak pola, langkah terakhir adalah interpretasi dan evaluasi pengetahuan (Alasadi & Bhaya, 2017).

2.2.3 *Machine Learning*

Machine learning adalah teknik yang menggunakan algoritma serta statistik komputasi untuk belajar dari data tanpa diprogram secara eksplisit. Teknik *machine learning* digunakan untuk menemukan pola dalam data yang kompleks. Pola serta pengetahuan tersembunyi dapat digunakan untuk memprediksi peristiwa pada masa depan dan melakukan pengambilan keputusan yang kompleks (Nikam, 2015).

Machine learning dikategorikan menjadi *supervised leaning*, *unsupervised learning* dan *reinforcement learning*. *Supervised learning* mengambil data berlabel dan membuat model untuk memprediksi dari data baru. Kategori ini dapat berupa masalah *clasification* dan *regression*. *Unsupervised learning* mengambil data tidak berlabel untuk menemukan pola dan membuat struktur dalam data.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Machine learning adalah program komputer yang dapat dilatih dengan data. Jenis data dalam *machine learning* yaitu data terstruktur dan data tidak terstruktur. Data terstruktur misalnya nama, umur dan tempat tanggal lahir. Data tersebut dapat disimpan dalam tabel berbasis baris dan kolom. Data tidak terstruktur misalnya file gambar, video dan suara. Data tersebut tidak dapat disimpan dalam tabel.

2.2.4 Klasifikasi

Teknik klasifikasi dalam *data mining* digunakan untuk memprediksi *class label* dengan tipe kategorikal dan mengklasifikasikan data berdasarkan kumpulan data *training* dan *class label* yang dapat digunakan untuk mengklasifikasikan data baru. Pemodelan klasifikasi dari sekumpulan data dengan *class label* yang telah diketahui disebut sebagai *pattern recognition* atau *supervised learning*. Klasifikasi meliputi *training* dan *testing*.

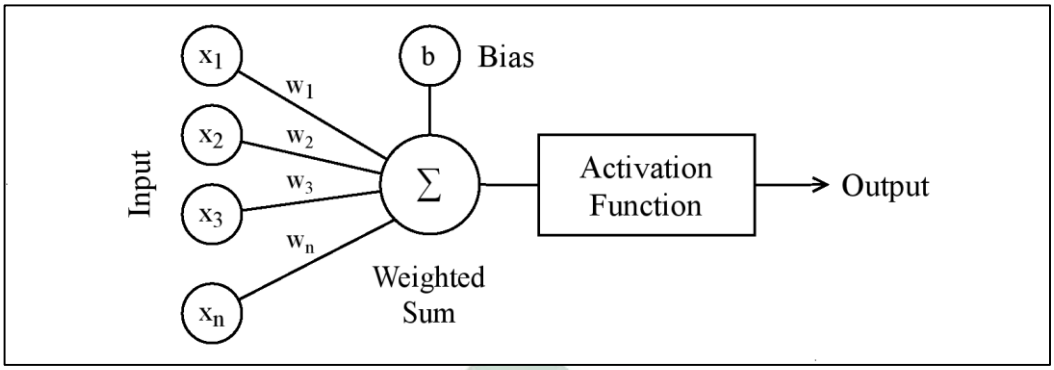
Training adalah proses menerapkan algoritma klasifikasi pada *training dataset*. Kemudian, *testing* adalah proses pengujian model terhadap *testing dataset* untuk mengukur kinerja dan akurasi model. Sehingga, dapat disimpulkan bahwa klasifikasi adalah proses untuk menetapkan *label* dari *dataset* yang belum diketahui atau tidak memiliki *class label*.

Algoritma dalam teknik klasifikasi yaitu *ID3*, *C4.5*, *K-Nearest Neighbors*, *Naïve Bayes*, *Support Vector Machine* dan *Artificial Neural Network*. Algoritma tersebut dapat digunakan dalam berbagai situasi sesuai kebutuhan. Selain itu, dapat diimplementasikan pada berbagai jenis *dataset*. Namun, setiap algoritma memiliki fitur dan keterbatasannya masing-masing (Nikam, 2015).

2.2.5 *Artificial Neural Network*

Artificial neural network (*ANN*) terinspirasi oleh jaringan saraf manusia. *Neural network* paling sederhana terdiri dari satu *neuron* dan disebut *perceptron*. *Perceptron* memiliki satu *input layer* dan satu *neuron*. *Input layer* bertindak sebagai dendrit yang berfungsi untuk menerima input. Jumlah *node* dalam *input layer* sama dengan jumlah *input dataset*. Gambar 2.1 menunjukkan diagram untuk menghitung *activation function* pada sebuah *neuron*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 2.1 *Single Perceptron Illustration* (Guna & Putri, 2020)

Activation function didapatkan dari jumlah dari perkalian nilai input dengan *weight*. *Weight* merupakan nilai acak yang telah diinisialisasi. *Activation function* berfungsi untuk memproses informasi dan menghasilkan output. Output merupakan hasil akhir pada *single perceptron*. Sedangkan pada *multilayer perceptron*, output dari *neuron* pada *layer* sebelumnya berfungsi sebagai *input* dari *neuron* pada *layer* selanjutnya (Heidari, Faris, Aljarah, & Mirjalili, 2018). Berikut adalah persamaan untuk menentukan nilai *weighted sum*.

$$z = w_1x_1 + w_2x_2 + \ldots + w_nx_n + b \tag{2.1}$$

$$z = \sum_{i=1}^{n} (w_ix_i) + b \tag{2.2}$$

Dalam persamaan tersebut, *w* adalah *weight* menunjukkan kekuatan node, *x* adalah *input* dari *neuron* dan *b* adalah *biases* yang menunjukkan nilai *weighted sum* yang harus dicapai *neuron* untuk aktif. Berdasarkan *weighted* sum, dapat diperoleh output dengan *activation functions*. Fungsi berupa *ReLu* untuk *hidden layer* dalam *multilayer perceptron*. *Softmax* untuk *output layer* dalam *multiclass classification*.

Activation function berfungsi untuk mempertahankan *output* dalam kisaran tertentu, sehingga membuat model efisien dan memiliki akurasi yang baik. *Sigmoid function* adalah fungsi logistik yang membantu menormalkan *output* dalam kisaran 0 hingga 1. Fungsi digunakan untuk klasifikasi *binary class*. Berikut persamaannya.

$$a = \frac{1}{1 + e^{-z}} \tag{2.3}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Softmax merupakan generalisasi dari fungsi *sigmoid* dan digunakan untuk klasifikasi *multiclass*. *Softmax* mendapatkan probabilitas masing-masing *class* yang jika dijumlahkan harus sama dengan 1. Ketika probabilitas suatu *class* meningkat, maka probabilitas *class* lain menurun. Sehingga, *class* dengan probabilitas tertinggi merupakan *output class*. Berikut adalah persamaan untuk fungsi *softmax*.

$$a_i = \frac{e_i^z}{\sum_{j=1}^{K} e_j^z} \tag{2.4}$$

Dalam persamaan tersebut, σ adalah fungsi *softmax*, z adalah nilai *weighted sum*, e^z adalah fungsi eksponensial standar untuk input z dan K adalah jumlah *class* dalam klasifikasi. Fungsi aktivasi untuk *hidden layer* meliputi *ReLu*, *Sigmoid* dan *Than*. Dalam *multilayer perceptron*, fungsi aktivasi menggunakan *ReLu*.

ReLu adalah *activation function* yang mengubah nilai *output* dalam kisaran 0 hingga *infinity*. *ReLu* akan melakukan konversi nilai negatif menjadi 0, sehingga tidak terdapat hasil negatif. Berikut adalah persamaan dari *ReLu function*.

$$a = \begin{matrix} \{z \; if \; z \geq 0 \\ 0 \; if \; z \leq 0\} \end{matrix} \tag{2.5}$$

Artificial neural network atau *multilayer perceptron* adalah kombinasi dari beberapa neuron yang terhubung dalam bentuk jaringan. *ANN* memiliki *input layer*, sejumlah *hidden layers* serta terakhir *output layer*. Pemodelan menggunakan *ANN* terdiri dari fase *Feed Forward* dan *Back Propagation*. Langkah yang dilakukan dalam fase *Feed Forward*, antara lain.

1. Nilai pada *input layer* diterima dan dikalikan dengan *weight*. *Bias* ditambahkan pada penjumlahan *weight*. Setiap *neuron* pada *hidden layer* pertama menerima nilai berbeda dari *input layer* bergantung pada *weight* dan bias.
2. *Neuron* memiliki *activation function* yang aktif berdasarkan nilai yang diterima dari *input*. Secara umum, *relu function* digunakan dalam *hidden layer neuron* dan *sigmoid function* digunakan untuk *output layer neuron*.
3. Output dari *hidden layer* pertama dikalikan dengan *weight* pada *hidden layer* berikutnya, hasilnya dijumlahkan dan diteruskan pada *layer* berikutnya. Proses ini berlangsung sampai *outer layer*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Parameter yang perlu ditetapkan yaitu jumlah *hidden layer*, jumlah *node* pada setiap *hidden layer*, jumlah iterasi atau *epoch*, *activation function* dan *cost optimizer*. Satu siklus yang meliputi *feed forward* dan *back propagation* disebut *epoch*. Langkah yang dilakukan dalam fase *Back Propagation*, antara lain.

1. *Loss function* digunakan untuk menghitung *error* dengan mengukur perbedaan antara *predicted output* dan *desired output*. *Loss function* yang digunakan dapat berupa *mean squared error* atau *cross entropy function*.
2. *Gradient descent* untuk *minimize error* dengan menghitung *partial derivation* dari *error function* dengan *weight* dan *bias*. *Gradient descent* digunakan untuk menemukan *local minimum* dari *error function* dengan mengubah nilai *weight* dan *bias* menggunakan nilai turunan parsial dari *error function*.

2.2.6 *Adam Optimizer*

Adam adalah algoritma yang mengoptimalkan fungsi *objective stochastic* berdasarkan perkiraan adaptif momen. Aturan pembaruan parameter *weights* dan *bias* dengan *Adam* adalah kombinasi dari metode optimasi *AdaGrad* dan *RMSProp*. *Adam* atau *Adaptive Moment Estimation* adalah metode yang efisien karena hanya membutuhkan nilai gradient urutan pertama. Fungsi gradient dapat diperoleh dari nilai turunan *loss function* atau *cross entropy*.

Algoritma menghitung *learning rate* adaptif individu untuk setiap parameter *weight* dan *bias* dari *first moment* dan *second moment estimation* dari gradien. *Adam* adalah metode yang menggabungkan *AdaGrad* yang bekerja dengan baik pada *sparse gradient* dan *RMSProp* yang bekerja yang baik dalam pengaturan *on-line* dan *non-stationary* (Kingman & Ba, 2015).

Adam menggunakan nilai β_1 dan β_2 untuk menghitung rata-rata bergerak. β_1 adalah *exponential decay of the rate* untuk perkiraan moment pertama dengan nilai default 0.9. β_2 adalah *exponential decay of the rate* untuk perkiraan moment kedua dengan nilai default 0.999. Berikut persamaan rata-rata bergerak.

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1) g_t \tag{2.6}$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2) {g_t}^2 \tag{2.7}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Dalam persaman tersebut, nilai dari rata-rata bergerak dihitung berdasarkan parameter β_1, β_2 dan gradient g_t untuk setiap iterasi *t*. Algoritma yang bergantung pada rata-rata bergerak sebagian besar mengalami *biases*, sehingga perlu fase untuk memperbaiki bias. Fase tersebut disebut dengan *bias correction step*.

$$\widehat{m}_t = \frac{m_t}{1 - \beta_1^t} \tag{2.8}$$

$$\hat{v}_t = \frac{v_t}{1 - \beta_2^t} \tag{2.9}$$

Berdasarkan persamaan tersebut, dapat dilakukan pembaruan nilai *weights* dan *biases* berdasarkan rata-rata bergerak terhitung dengan nilai *step size* η. Berikut adalah persamaan untuk memperbarui parameter *weight* dan *bias*.

$$w_t = w_{t-1} - \eta \frac{\widehat{m}_t}{\sqrt{\hat{v}_t} + \varepsilon} \tag{2.10}$$

Parameter *weights* dan *biases* diperbarui dengan memberikan nilai *step size*. *Step size* memiliki nilai default sebesar $\eta = 0.001$ dan ε epsilon adalah nilai untuk mencegah pembagian nol dan memiliki nilai default sebesar $\varepsilon = 10^{-8}$.

Algoritma *adam* memiliki kecepatan pelatihan lebih baik daripada metode optimasi lainnya seperti *AdaGrad*, *RMSProp*, *SGD*, *Nesterov* dan *AdaDelta* (Bock, Goppold, & Wei, 2018). *Adam* juga memiliki kinerja lebih baik daripada metode optimasi lainnya pada iterasi awal (Huda et al., 2019).

2.2.7 Kombinasi *Adam* dan *Nesterov*

Nesterov-accelerated Adam (*Nadam*) adalah algoritma yang mengabungkan *Adam* dengan *Nesterov Momentum*. Modifikasi komponen momentum pada *Adam* dengan memanfaatkan *Nesterov's Accelerated Gradient* (*NAG*) dilakukan untuk meningkatkan kecepatan dan kualitas *Neural Network* (Dozat, 2016). Pembelajaran dengan metode *Nadam* dipercepat dengan meringkas nilai *exponential decay* rata-rata bergerak dari gradien sebelumnya dan gradien saat ini.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Modifikasi nilai inisialisai dilakukan dengan satu kali untuk memperbarui gradien dan kedua kalinya untuk memperbarui momentum, kemudian menerapkan vektor momentum *look-ahead* untuk memperbarui *weight*.

$$g_t = \nabla_{w_t} J(w_t) \tag{2.11}$$

$$m_t = \gamma m_{t-1} + \eta g_t \tag{2.12}$$

$$w_{t-1} = w_t - (\gamma m_t + \eta g_t) \tag{2.13}$$

Dalam persamaan tersebut, J adalah nilai loss function, γ adalah momentum decay dan η merupakan ukuran langkah. Modifikasi metode *adam* hanya dilakukan pada $\widehat{m}_t$ dan tidak pada $\widehat{v}_t$. Berikut persamaan untuk memperbarui *weight*.

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{\widehat{v}_t} + \epsilon}(\beta_1 \widehat{m}_{t-1} + \frac{(1+\beta_1) g_t}{1-\beta_1^t}) \tag{2.14}$$

2.2.8 *Cross Entropy*

Cross entropy atau *log loss* merupakan *loss function* yang berfungsi untuk mengukur kinerja dari model klasifikasi. *Log loss* akan meningkat jika probabilitas prediksi semakin menyimpang dari *actual clas*. Semakin rendah nilai *cross entropy*, maka semakin baik model klasifikasi. Hasil terbaik untuk *cross entropy* adalah 0. *Loss function* untuk *binary classification* dapat dihitung sebagai berikut.

$$L = -(y \log(p) + (1-y) \log(1-p)) \tag{2.15}$$

Cross entropy untuk *multiclass classification* digunakan untuk model yang memiliki *class label* lebih dari dua. *Multiclass cross entropy* menghitung nilai *loss* secara terpisah untuk setiap *class label* pada setiap observasi dan menjumlahkan nilai *loss* tersebut. Berikut adalah persamaan dari *multiclass cross entropy*.

$$L = -\sum_{c=1}^{M} y_c \log(p_c) \tag{2.16}$$

Dalam persamaan tersebut, *M* adalah jumlah *class labels* dalam klasifikasi, *y* merupakan nilai probabilitas aktual dan *p* adalah nilai probabilitas yang diprediksi untuk setiap observasi prediksi dari *class c*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.2.9 *Confusion Matrix*

Confusion matrix yakni teknik yang berfungsi untuk mendapatkan kinerja model klasifikasi dan menggambarkan hasil prediksi dari klasifikasi. Hasil prediksi yang benar dan salah didistribusikan dengan nilai hitungan untuk setiap *class label*. *Confusion matrix* mengungkapkan angka *true positive*, *false positive*, *true negative* dan *false negative* (Khan & Rana, 2019).

Tabel 2.2 Ilustrasi *Confusion Matrix*

Predicted Value	***Actual Value***	
	Positive Class	*Negative Class*
Positive Class	*True Positive* (*TP*)	*False Positive* (*FP*)
Negative Class	*False Negative* (*FN*)	*True Negative* (*TN*)

Berdasarkan hasil *confusion matrix* dapat diketahui *accurary*, *precision* dan *recall* pada sebuah model klasifikasi. *Accuracy* menggambarkan nilai perbandingan jumlah prediksi benar dan jumlah semua data masukan. Berikut adalah persamaan untuk menghitung nilai *accuracy*.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2.17}$$

Precision adalah nilai yang mendeskripsikan tingkat kesamaan antara data yang diminta terhadap hasil klasifikasi. Re*call* mendeskripsikan kesuksesan model dalam mendapatkan ulang informasi. Berikut adalah persamaan untuk mengetahui nilai *accuracy*, *precision* dan *recall* dari model klasifikasi.

$$Precision = \frac{TP}{TP+FP} \tag{2.18}$$

$$Recall = \frac{TP}{TP+TN} \tag{2.19}$$

Klasifikasi multiclass tidak memiliki kelas positif dan negatif seperti *binary classification*, sehingga untuk menemukan angka *true positive*, *false positive*, *true negative* dan *false negative* perlu dilakukan perhitungan pada setiap *class*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.2.10 *Python Programming Language*

Python adalah bahasa pemrograman *Machine Learning* dan *Deep Learning* yang dilengkapi dengan *library* untuk menyelesaikan *use case* dan *project*. Bahasa ini merupakan bahasa program tingkat tinggi yang *cross platform* dan *readable code*. *Python libraries* untuk kebutuhan *machine learning* yaitu *Tensorflow*, *Keras*, *Scikit-learn*, *Pandas*, *Numpy* dan lain sebagainya (Blank & Deb, 2020).

Tensorflow memiliki fungsi inti untuk membangun model *Deep Learning*. *Keras* adalah *library* yang dapat diakses dalam *tensorflow*. Manipulasi dan analisis data dilakukan oleh *Pandas*. *Library* ini melakukan persiapan data sebelum proses *training* dan menghasilkan data multidimensi terstruktur. *Pandas* memiliki fitur untuk melakukan *reshape*, *pivot*, *merge* dan *joining* dataset. Selain itu, *library* ini juga menangani *missing value*, *alignment*, *indexing* dan *filtering*.

2.2.11 *Scikit Learn Library*

Scikit-learn adalah alat sederhana dan efisien untuk analisis data prediktif. *Library* ini dibangun dengan *Numpy*, *Scipy* dan *matplotlib*. *Scikit-learn* mendukung beberapa algoritma untuk melakukan klasifikasi, regresi, *clustering*, *dimensionaly reduction*, *model selection* dan *preprocessing*. Algoritma klasifikasi yang didukung oleh *scikit-learn* yaitu *nearest neighbors*, *naïve bayes*, *decision tree*, *neural network* dan lain-lain. Scikit-*learn* berfungsi untuk pemodelan data.

2.2.12 *Flask Framework*

Flask adalah *micro-framework* untuk pengembangan *website* yang ditulis dalam bahasa pemrograman *python*. Kerangka kerja dikembangkan untuk membuat aplikasi *python* dan juga membantu proses implementasi aplikasi *machine* learning sebagai aplikasi *web*. *Flask* memiliki dua komponen utama yaitu *WSGI toolkit* dan *Jinja2 template engine*. *WSGI* adalah spesifikasi untuk aplikasi *website* dan *Jinja2* untuk membuat halaman *website*.

Flask memiliki fitur yang sederhana, ringan, tidak bergantung pada *library* tertentu dan menerapkan *bare minimum webserver*, sehingga disebut dengan *micro-framework*. Secara umum, *flask* menyediakan *Wekzeug* untuk menerima *request* dan *response* (Mufid, Rasyid, Rochimansyah, & Rokhim, 2019).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

2.3 Integrasi Keilmuan

Berdasarkan hasil wawancara kepada pakar yaitu ibu Wiwin Luqna Hunaida dosen Pendidikan Agama Islam Universitas Islam Negeri Sunan Ampel, terdapat hadis yang memiliki keterkaitan terhadap kesiapan sekolah dan syarat umur anak untuk masuk sekolah dasar. Hadis tersebut tentang perintah salat kepada anak sejak umur tujuh tahun yang diriwayatkan oleh Abu Daud sebagai berikut.

حَدَّثَنَا مُحَمَّدُ بْنُ عِيسَى يَعْنِي ابْنَ الطَّبَّاعِ حَدَّثَنَا إِبْرَاهِيمُ بْنُ سَعْدٍ عَنْ عَبْدِ الْمَلِكِ بْنِ الرَّبِيعِ بْنِ سَبْرَةَ عَنْ أَبِيهِ عَنْ جَدِّهِ قَالَ قَالَ النَّبِيُّ صَلَّى اللَّهُ عَلَيْهِ وَسَلَّمَ مُرُوا الصَّبِيَّ بِالصَّلَاةِ إِذَا بَلَغَ سَبْعَ سِنِينَ وَإِذَا بَلَغَ عَشْرَ سِنِينَ فَاضْرِبُوهُ عَلَيْهَا

Artinya: Telah menceritakan kepada kami Muhammad bin Isa bin Ali bin Abi Thalib-Thabba' dari Ibrahim bin Sa'd dari Abdul Malik bin Ar-Rabi' bin Sabrah dari Ayahnya dari Kakeknya dia berkata; Nabi shallallahu 'alaihi wasallam bersabda: "Perintahkan anak-anakmu untuk menjalankan salat jika telah memasuki umur tujuh tahun dan jika mereka telah memasuki umur sepuluh tahun, kamu dapat memukul dirinya bilamana tidak melaksanakan salat" (HR. Abu Daud).

Hadis tersebut menjelaskan bahwa orang tua wajib mengenalkan dan juga memerintah anak untuk salat sewaktu anak sudah menginjak usia tujuh tahun. Pada usia tersebut, diri anak sedang menduduki usia *mumayyiz*, ketika anak telah dapat membedakan kiri dan kanan (Al-Jauziyah, 1990).

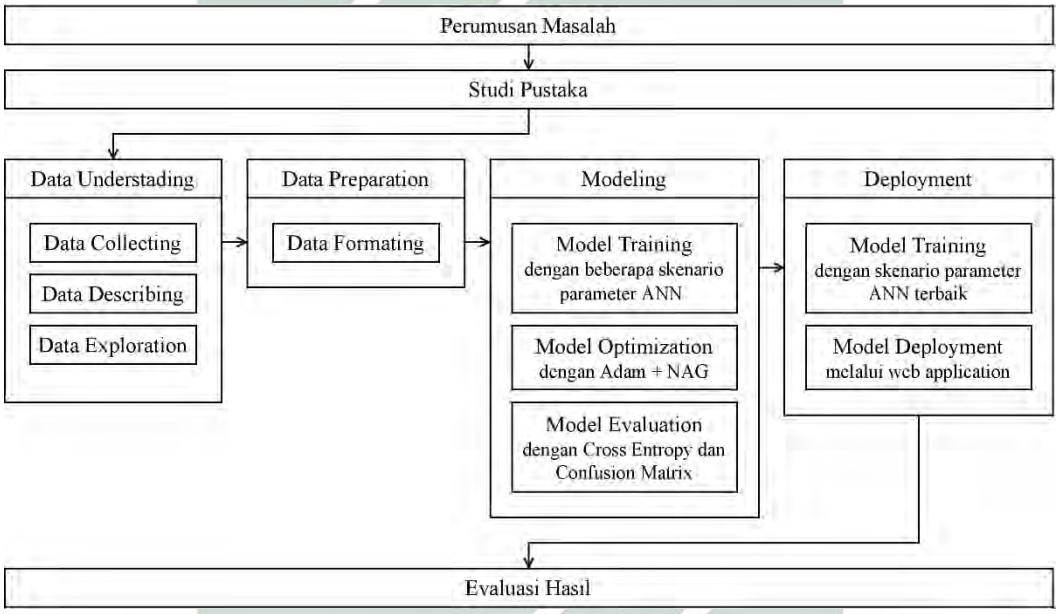
Hal tersebut memiliki keterkaitan terhadap syarat masuk sekolah dasar yang juga menyebutkan bahwa usia yang tepat untuk masuk sekolah dasar adalah tujuh tahun. Usia tersebut dinilai tepat, karena anak telah memiliki kesiapan fisik juga dan psikis untuk mengikuti pendidikan formal. Usia dapat dipandang sebagai suatu variable yang berdampak pada kesiapan sekolah. Sehingga, usia dapat digunakan untuk memprediksi kesiapan anak untuk sekolah.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB III
METODOLOGI PENELITIAN

3.1 Metode Penelitian

Metode penelitian menjelaskan langkah-langkah yang akan dilakukan untuk mendapatkan hasil penelitian. Diagram alur digunakan dalam penguraian metode penelitian. Gambar 3.1 menunjukkan metode dalam penelitian ini.



Gambar 3.1 Metode Penelitian

Hasil penelitian didapatkan dari proses perumusan masalah, studi pustaka, *data understanding*, *data preparation*, *modeling*, *deployment* dan evaluasi hasil. Berikut adalah penjelasan terkait setiap proses dalam penelitian.

3.1.1 Perumusan Masalah

Perumusan masalah merupakan tahap inisialisai penelitian, sehingga dapat diketahui masalah yang akan diselesaikan dan hasil yang akan dicapai. Berdasarkan perumusan masalah, penelitian bertujuan untuk mengetahui kinerja dan akurasi *Artificial Neural Network* dengan metode kombinasi *Adam* dan *Nesterov* dalam melakukan klasifikasi dan prediksi kesiapan sekolah.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

3.1.2 Studi Pustaka

Studi pustaka adalah proses untuk merangkum penelitian sebelumnya yang memiliki relasi dengan penelitian ini. Hasil studi pustaka dapat digunakan sebagai referensi dan pengetahuan untuk mengembangkan penelitian. Berdasarkan studi pustaka, *Artificial Neural Network* memiliki akurasi lebih baik daripada algoritma lainnya dalam memprediksi kesiapan sekolah. Tinjauan pustaka juga menunjukkan, *Adam* dan *Nesterov* dapat digunakan untuk meningkatkan kinerja model *ANN*.

3.1.3 *Data Understanding*

Data understanding meliputi proses *data collecting*, *data describing* dan *data exploration*. *Data collecting* adalah proses mengumpulkan data administratif *Nijmeegse Schoolbekwaamheids Test* yang diperoleh dari TK Ar-Rasyid Surabaya. *Data describing* adalah proses menyeleksi dan mendekripsikan *attribute* dalam data administratif *NST* yang memiliki pengaruh terhadap kesiapan sekolah.

Pemilihan *attribute* untuk memprediksi kesiapan sekolah dapat didasarkan pada penelitian terdahulu. Tabel 3.1 menunjukkan kumpulan *data attribute*.

Tabel 3.1 Atribut Penelitian Terdahulu

No	Data Atribut	(Suleiman et al., 2017)	(Nurlina et al., 2019)	(Intan et al., 2019)
1	*Age*		✓	✓
2	*Gender*		✓	✓
3	*Child Position*	✓	✓	✓
4	*Number of Siblings*		✓	✓
5	*Family Size*	✓		
6	*Have Pre-Elementary*		✓	✓
7	*Father Education*	✓	✓	✓
8	*Mother Education*	✓	✓	✓
9	*Father Occupation*	✓		
10	*Mother Occupation*	✓		
11	*Neighborhood Peers*	✓		
12	*Class Peers*	✓		

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berdasarkan penelitian terdahulu, atribut yang digunakan dalam penelitian ini yaitu age, gender, child position, number of siblings, have pre-elementary, father education, mother education dan school readiness.

Tabel 3.2 Deskripsi Fitur dan Label

No	Data Atribut	Tipe Data	Deskripsi
1	*Age*	*Numeric*	Age digunakan untuk klasifikasi pada segmentasi umur. Contoh: 6 Tahun, 7 Tahun.
2	*Gender*	*Categorical*	Gender digunakan untuk klasifikasi pada segmentasi jenis kelamin. Contoh: Laki-laki, Perempuan.
3	*Child position*	*Numeric*	Child position digunakan untuk klasifikasi segmentasi urutan anak. Contoh: Pertama, Ketiga.
4	*Number of siblings*	*Numeric*	Atribut digunakan untuk klasifikasi pada segmentasi jumlah saudara. Contoh: 1, 3.
5	*Have pre-elementary*	*Categorical*	Atribut digunakan untuk klasifikasi pada segmentasi status prasekolah. Contoh: Pernah, Tidak Pernah.
6	*Father education*	*Categorical*	Atribut digunakan untuk klasifikasi pada segmentasi pendidikan ayah. Contoh: SMA, S1.
7	*Mother education*	*Categorical*	Atribut digunakan untuk klasifikasi pada segmentasi pendidikan ibu. Contoh: SMA, S1.
8	*School Readiness (Class Label)*	*Categorical*	Atribut yang akan diprediksi. *Class*: Dipertimbangkan, Disarankan dan Tidak Disarankan.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Data exploration adalah proses mengungkapkan ringkasan menarik terkait hasil tes kesiapan sekolah. Data administratif *Nijmeegse Schoolbekwaamheids Test* digunakan untuk menampilkan distribusi hasil tes kesiapan sekolah terhadap setiap atribut yaitu umur, jenis kelamin, urutan anak, jumlah saudara, status prasekolah, pendidikan ayah dan pendidikan ibu. Berdasarkan grafik distribusi, dapat diketahui korelasi antara hasil tes kesiapan sekolah dengan setiap atribut.

3.1.4 *Data Preparation*

Data preparation mencakup proses *data formating*. *Data formating* adalah proses yang berfungsi untuk memudahkan model *ANN* dalam melakukan proses klasifikasi dengan mengubah format data. Contohnya nilai dalam atribut *have pre-elementary*, "Pernah" dapat diubah dengan nilai 1 dan "Tidak Pernah" dengan nilai 0. *Data formating* dilakukan dengan menggunakan *Pandas* dan *Scikit-learn*.

3.1.5 *Modeling*

Modeling meliputi proses *training*, *optimization* dan *evaluation*. *Training* dilakukan untuk menemukan pola data. *Artificial neural network* dengan beberapa skenario parameter digunakan dalam proses *training*. Hal tersebut dilakukan untuk menemukan arsitektur *ANN* dengan kinerja terbaik.

Optimization dilakukan menggunakan *Adam Optimizer*. Algoritma tersebut diharapkan dapat meningkatkan kinerja *ANN*. *Adam* juga akan dimodifikasi dengan memanfaatkan *Nesterov* (*NAG*). *Evaluation* dilakukan menggunakan *cross entropy* dan *confusion matrix* untuk mengetahui kinerja *ANN*. Proses *training*, *optimization* dan *evaluation* dilakukan dengan *Scikit-learn*.

Tabel 3.3 Parameter Model *Artificial Neural Network*

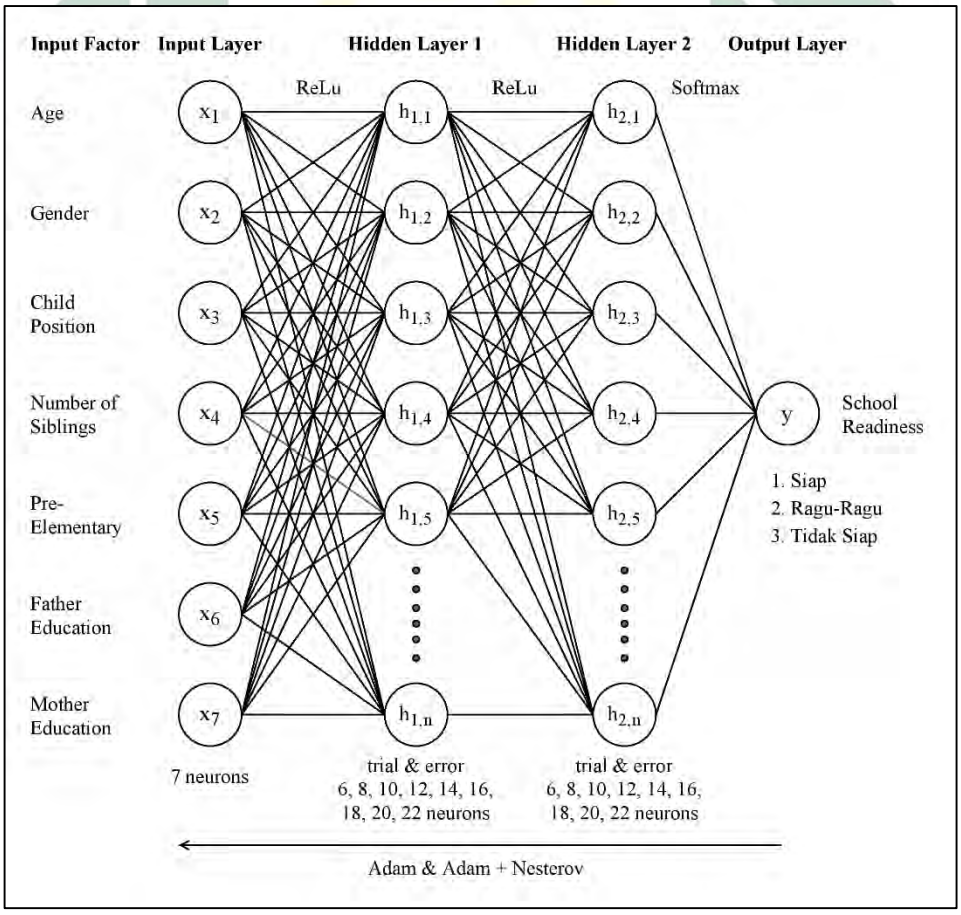
No	Parameter	Nilai
1	*Number of Input*	7
2	*Number of Layes*	2
3	*Number of Neuron*	6, 8, 10, 12, 14, 16, 18, 20, 22 (trial and error)
4	*Number of Output*	3
5	*Hidden Layer Activation*	ReLu

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 3.4 Parameter Model *Artificial Neural Network* lanjutan

No	Parameter	Nilai
6	*Output Layer Activation*	Softmax
7	*Optimizer*	Adam, Adam + Nesterov (trial and error)
8	*Loss Function*	Cross Entropy
9	*Batch Size*	32
10	*Epoch*	500, 1000 (trial and error)

Arsitektur *ANN* terdiri dari *input layer* dengan 7 buah neuron, 2 *hidden layer* dengan menguji sejumlah neuron serta *output layer* dengan 1 neuron. *Output layer* menggambarkan hal yang diprediksi, yaitu *School Readiness*. Hasil prediksi akan mengklasifikasikan anak dalam kelompok Dipertimbangkan, Disarankan dan Tidak Disarankan untuk masuk sekolah dasar. Arsitektur *ANN* ditunjukkan Gambar 3.2.



Gambar 3.2 Arsitektur Model *Artificial Neural Network*

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

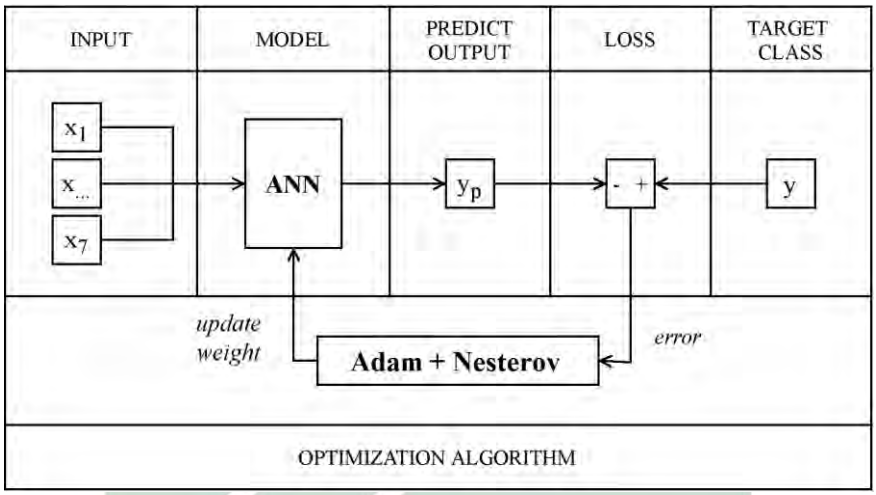
Berdasarkan parameter model *ANN* yang telah ditentukan pada Tabel 3.3, terdapat 36 skenario yang digunakan dalam proses *training*. *Training* dilakukan berulang kali dengan mengubah nilai dari parameter *number of neurons*, *optimizer* dan *number of epoch* untuk menghasilkan model *ANN* dengan kinerja terbaik. Tabel 3.5 menunjukkan skenario perlatihan *ANN*.

Tabel 3.5 Skenario Pelatihan Model *Artificial Neural Network*

Skenario	***Neuron***	***Epoch***	***Optimizer***
1	6	500	Adam
2	6	1000	Adam
3	6	500	Adam + Nesterov
4	6	1000	Adam + Nesterov
5	8	500	Adam
6	8	1000	Adam
7	8	500	Adam + Nesterov
8	8	1000	Adam + Nesterov
…	…	…	…
30	20	1000	Adam
31	20	500	Adam + Nesterov
32	20	1000	Adam + Nesterov
33	22	500	Adam
34	22	1000	Adam
35	22	500	Adam + Nesterov
36	22	1000	Adam + Nesterov

Pengujian model dilakukan menggunakan skenario pembagian data 5-*fold cross validation*. Skenario ini melakukan *training* dan *testing* dengan membagi data dalam lima bagian dengan empat bagian untuk data *training* dan satu bagian untuk data *testing*. Proses pelatihan dan pengujian dilakukan sebanyak lima kali sampai seluruh bagian data telah menjadi *training* dan *testing*. Berikut diagram block untuk menunjukkan fungsi *Adam* dan *Nesterov* dalam *Artificial Neural Network*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 3.3 Diagram Block Model *Artificial Neural Network*

Berdasarkan diagram, kombinasi *Adam* dan *Nesterov Accelerated Gradient* menggunakan nilai *error* untuk memperbarui nilai *weight* dari model *ANN*. Proses tersebut dilakukan hingga mencapai minimal *loss*.

3.1.6 *Deployment*

Deployment meliputi proses penyimpanan dan penerapan model *Artificial Neural Network* untuk memprediksi kesiapan sekolah. Penyimpanan model *ANN* dilakukan menggunakan fungsi *dump model* pada *pickle*. Model *ANN* yang telah tersimpan dalam format file tertentu, kemudian diterapkan dalam pengembangan sistem prediksi menggunakan fungsi *load model* pada *pickle*. Pengembangan sistem prediksi kesiapan sekolah dilakukan menggunakan *Flask Framework*.

3.1.7 Evaluasi Hasil

Evaluasi hasil adalah proses menguraikan dan merangkum hasil penelitian terkait dengan kinerja model *Artificial Neural Network* dengan metode kombinasi *Adam* dan *Nesterov* dalam melakukan prediksi kesiapan sekolah.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB IV
HASIL DAN PEMBAHASAN

4.1 *Data Understanding*

Data penelitian berasal dari data administratif anak yang telah mengikuti tes kesiapan masuk Sekolah Dasar berbasis *Nijmeegse Schoolbekwaamheids Test* 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.

Descriptive statistics untuk categorical data terdiri dari nilai frequency dan percentage. Statistik tersebut digunakan untuk mengetahui persebaran data setiap variable. Tabel 4.1 menunjukkan persebaran data dari variable jenis kelamin, status prasekolah, pendidikan ayah dan pendidikan ibu.

Tabel 4.1 Descriptive Statistics Categorical Features

Variable	***Value***	***Frequency***	***Percentage***
Gender	Laki-laki	124	55.11 %
	Perempuan	101	44.89 %
Have Pre-Elementary	Pernah	156	69.33 %
	Tidak Pernah	69	30.67 %
Father Education	SMP	2	0.89 %
	SMA	57	25.33 %
	Diploma	5	2.22 %
	Sarjana	161	71.56 %
Mother Education	SMP	2	0.89 %
	SMA	54	24.00 %
	Diploma	2	0.89 %
	Sarjana	167	74.22 %

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Hasil persebaran data menunjukkan jumlah laki-laki dan perempuan yang mengikuti tes kesiapan sekolah tidak jauh berbeda. Namun, terdapat perbedaan jauh pada persebaran pendidikan orang tua. Berdasarkan data penelitian yang diperoleh sebanyak 225 anak, sekitar 70% orang tua memiliki pendidikan Sarjana dan sisanya memiliki pendidikan SMP, SMA dan Diploma.

Statistik deskriptif untuk continuous variable meliputi nilai mean, standard deviation, minimum dan maximum. Nilai standard deviasi digunakan untuk melihat variasi dari data. Tabel 4.2 menunjukkan persebaran data dari variable umur, urutan anak dan jumlah saudara.

Tabel 4.2 Descriptive Statistics Numerical Features

Variable	***Mean***	***Std Dev***	***Minimum***	***Maximum***
Age	5.79	0.82	5	7
Child Position	1.91	0.79	1	5
Number of Siblings	1.51	0.91	0	4

Persebaran data menunjukkan nilai minimum dan maximum untuk seluruh variable tidak memiliki jangkauan yang besar. Nilai standard deviasi juga tidak jauh berbeda pada seluruh variable yaitu 0.79 - 0.91. Hal tersebut menunjukkan bahwa nilai umur, urutan anak dan jumlah saudara tidak terlalu bervariasi.

Hasil tes kesiapan sekolah dalam penelitian berfungsi sebagai variable target atau *class* yang akan diprediksi, sedangkan variable lainnya berfungsi sebagai fitur. Berdasarkan data yang telah diperoleh, tes kesiapan sekolah berbasis *NST* akan menunjukkan bahwa anak Disaranakan, Dipertimbangkan atau Tidak Disarankan untuk masuk jenjang pendidikan Sekolah Dasar.

Tabel 4.3 Jumlah Data Setiap Target Class

Class	***Frequency***	***Percentage***
Disarankan	133	59.11 %
Dipertimbangkan	27	12.00 %
Tidak Disarankan	65	28.89 %

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Penelitian ini menggunakan Chi-Square Test pada SPSS untuk mengetahui korelasi antara umur, jenis kelamin, urutan anak, jumlah saudara, status prasekolah, pendidikan ayah dan pendidikan ibu dengan kesiapan sekolah anak. Chi-Square test menggunakan cross tabulasi untuk menganalisis data. Hasil korelasi menggunakan Chi-Square Test ditunjukkan pada Tabel 4.4.

Tabel 4.4 Hubungan Variable Fitur dengan Variable Target

Variable		***Chi-Square***	***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 pre-elementary	↔ *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 correlation pada Tabel 4.4, terdapat hubungan significant yang kuat antara status prasekolah dan kesiapan sekolah dengan nilai chi^2=187.16 dan p<.05. Jumlah saudara dan kesiapan sekolah juga memiliki hubungan significan yang kuat dengan nilai chi^2=128.54 dan p<.05. Hubungan significant yang kuat juga terjadi antara umur dan kesiapan sekolah dengan nilai chi^2=102.32, p<.05.

Namun, terdapat hubungan significant yang lemah antara pendidikan ayah, pendidikan ibu, urutan anak dengan kesiapan sekolah dengan nilai chi^2 20.50-37.37 dan p<.05. Jenis kelamin dan kesiapan sekolah tidak memiliki hubungan significant dengan nilai chi^2=2.11 dan p>.05.

4.2 *Data Preparation*

Data administrastif *NST* memiliki beberapa atribut dengan tipe data *numeric* atau *categorical*. Data formating melakukan perubahan data *categorical* menjadi *numeric* untuk memudahkan pembangunan model *Artificial Neural Network* dalam melakukan klasifikasi. Atribute dengan tipe *categorical* yaitu jenis kelamin, status prasekolah, pendidikan ayah, pendidikan ibu dan kesiapan sekolah.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.5 Data Formating pada Data Categorical

No	Data Atribut	Nilai	Format
1	*Gender*	Laki-laki	0
		Perempuan	1
2	*Have Pre-Elementary*	Tidak Pernah	0
		Pernah	1
3	*Father Education & Mother Education*	SMP	0
		SMA	1
		Diploma	2
		Sarjana	3
4	*School Readiness*	Tidak Disarankan	0
		Dipertimbangkan	1
		Disarankan	2

Berdasarkan tabel format, pada atribut status prasekolah, pendidikan orang tua dan kesiapan sekolah menunjukkan bahwa semakin besar angka semakin positif nilai yang dimiliki. Hal tersebut menunjukkan bahwa perubahan nilai *categorical* menjadi *numeric* perlu mempertimbangkan tingkatan nilai dalam data.

4.3 *Modeling*

Penelitian menggunakan algoritma Artificial Neural Network. Algoritma ini menerapkan metode optimization untuk meningkatkan kinerja, dalam penelitian ini *Adam* dan *Nesterov* digunakan untuk memprediksi kesiapan sekolah. Perbedaan utama terletak pada komponen momentum yang digunakan. *Adam* menggunakan *classic momentum*, sedangkan *Adam Nesterov* menggunakan *nesterov momentum*.

Momentum method merupakan teknik yang dapat mempercepat penurunan gradien dengan memperhitungkan gradient sebelum dalam aturan pembaruan pada setiap iterasi. Perbedaan antara momentum dan nesterov accelerated gradien adalah fase komputasi gradien. Vanilla momentum, gradien dihitung menggunakan current parameter, sedangkan Nesterov menerapkan kecepatan ke parameter untuk mhitung interim parameters, kemudian digunakan untuk menghitung gradient.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Nesterov Accelerate Gradient dapat dipertimbangkan sebagai faktor koreksi untuk classical Momentum. Saat learning rate besar, Nesterov memungkinkan nilai *decayrate* yang lebih besar daripada metode Momentum, sehingga nesterov mampu mencegah terjadi oscillation atau repetitive variations. Osilasi dapat memperlambat model untuk mencapai *minimum loss*.

4.3.1 *Model Training*

Nesterov Accelerate Gradient diharapkan dapat meningkatkan performance Adam dalam memprediksi kesiapan sekolah. Adam memulai pelatihan data dengan melakukan inisialisasi step size, exponential decay of rates dan epsilon. Pseudocode berikut mendeskripsikan algoritma Adam dalam melakukan pembaruan weights.

```
SET stepSize to 0.001
SET firstDecay to 0.9
SET secondDecay to 0.999
SET epsilon to 10^-8
GET lossFunction
GET weighParameter
INIT firstMoment to 0
INIT secondMoment to 0
INIT timeStep to 0
WHILE weightParameter isnot converged
      INCREMENT timeStep
      GET gradient
      UPDATE firstMoment
      UPDATE secondMoment
      COMPUTE fistMomentCorrected
      COMPUTE secondMomentCorrected
      UPDATE weightParameter
ENDWHILE
RETURN weightParameter
```


digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Nesterov's Accelerated Gradient akan diterapkan Adam dengan melakukan modifikasi komponen momentum. Nesterov menggunakan current parameter untuk mengetahui nilai future loss. Jika pembaruan menyebabkan bad loss, maka gradien akan mengarahkan pembaruan kembali ke parameter sebelumnya.

```
SET stepSize to 0.001
SET epsilon to 10^-8
SET velocity
GET lossFunction
GET weighParameter
INIT firstMoment to 0
INIT secondMoment to 0

WHILE weightParameter isnot converged
      CALCULATE gradient W.R.T lossFunction
      UPDATE firstMoment W.R.T gradient
      UPDATE secondMoment W.R.T gradientsquare
      COMPUTE fistMomentCorrected
      COMPUTE secondMomentCorrected
      UPDATE newWeightParameter W.R.T currentWeightParameter
ENDWHILE

RETURN weightParameter
```

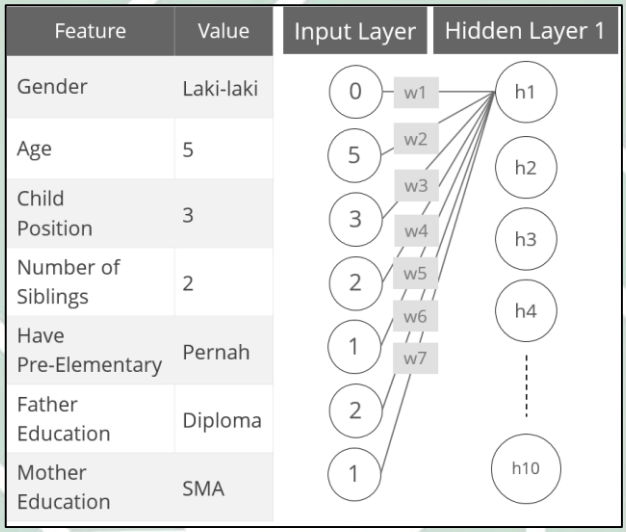
Step size memiliki nilai default sebesar $\eta = 0.001$ dan ε epsilon adalah nilai untuk mencegah pembagian nol dan memiliki nilai default $\varepsilon = 10^{-8}$. Perhitungan fungsi gradien diperoleh dari nilai turunan loss. Perkiraan moment pertama dihitung menggunakan nilai gradient dan yang kedua menggunakan gradient square.

Perkiraan moment pertama dan kedua selanjutnya melalui perhitungan bias correction step. Pembaruan parameter weight dilakukan berdasarkan curent weight. Proses tersebut dilakukan berulang sampai parameter weight converged.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Untuk membandingkan kinerja Adam dan metode kombinasi Adam dengan Nesterov Accelerated Gradient, penelitian ini menggunakan Scikit-Learn dan Grid Search untuk memudahkan pengujian. Namun, perhitungan manual tetap dilakukan untuk membangun pemahaman terkait algoritma Artificial Neural Network beserta metode optimization Adam dan kombinasi Adam dan Nesterov.

Perhitungan menggunakan sample dari data anak dengan jenis kelamin laki-laki, umur 5, anak ke-3, jumlah saudara 2, pernah mengikuti pendidikan prasekolah, pendidikan ayah adalah Diploma dan pendidikan ibu adalah SMA.



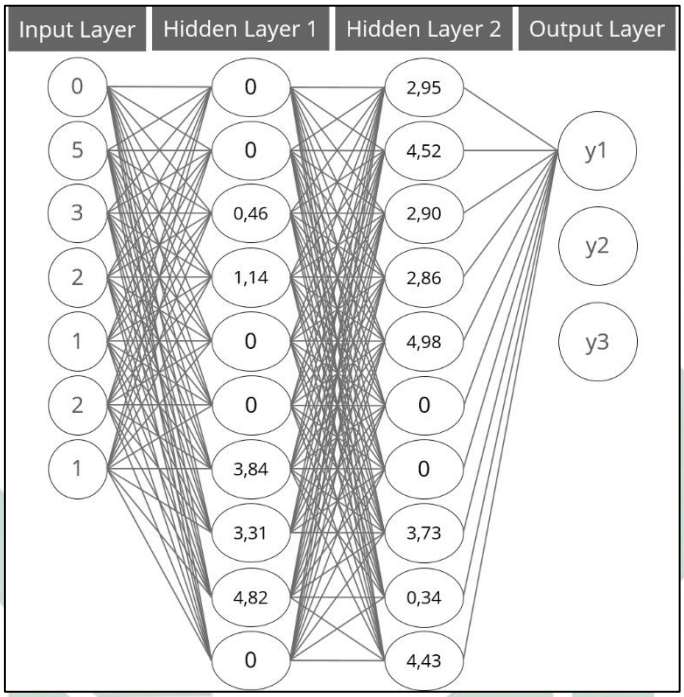
Gambar 4. 1 Hasil Perhitungan Neuron Pada Input Layer

Perhitungan neuron pada hidden layer pertama dilakukan dengan perkalian nilai input dan nilai weights. Nilai weight adalah nilai acak yang telah diinisialisasi. Untuk memudahkan perhitungan all neuron, proses perkalian menggunakan konsep perkalian matrix. Pehitungan menggunakan persamaan 2.2.

$$
\begin{vmatrix}
1{,}03 & -0{,}18 & 0{,}42 & 0{,}79 & -2{,}13 & -0{,}54 & -0{,}11 \\
0{,}42 & -0{,}49 & 0{,}25 & -0{,}38 & 0{,}55 & -0{,}14 & -0{,}28 \\
-0{,}39 & -0{,}45 & 0{,}29 & 0{,}87 & -0{,}09 & 0{,}07 & 0{,}31 \\
-0{,}33 & 0{,}49 & -0{,}34 & -0{,}77 & -0{,}07 & 0{,}41 & -0{,}1 \\
0{,}87 & 0{,}4 & -0{,}79 & 0{,}67 & -0{,}08 & -0{,}63 & -0{,}2 \\
-0{,}05 & -0{,}36 & -0{,}42 & -0{,}28 & -0{,}02 & -0{,}22 & 0{,}58 \\
-0{,}67 & 0{,}22 & 1{,}02 & -0{,}19 & -0{,}48 & -0{,}19 & 0{,}35 \\
0{,}75 & 0{,}68 & -0{,}43 & 0{,}62 & 1{,}04 & -0{,}3 & 0{,}21 \\
-0{,}99 & -0{,}15 & 0{,}57 & 1{,}06 & -0{,}35 & 0{,}99 & 0{,}32 \\
-0{,}53 & -0{,}53 & 0{,}06 & -0{,}44 & -0{,}42 & 0{,}39 & -0{,}52
\end{vmatrix}
\begin{vmatrix}
0 \\ 5 \\ 3 \\ 2 \\ 1 \\ 2 \\ 1
\end{vmatrix}
+
\begin{vmatrix}
0{,}49 \\ 0 \\ -0{,}26 \\ 0{,}6 \\ 0{,}32 \\ 0 \\ 0{,}57 \\ -0{,}69 \\ -0{,}21 \\ 0
\end{vmatrix}
=
\begin{vmatrix}
-0{,}89 \\ -2{,}47 \\ 0{,}46 \\ 1{,}14 \\ -0{,}25 \\ -3{,}5 \\ 3{,}84 \\ 3{,}31 \\ 4{,}82 \\ -3{,}51
\end{vmatrix}
$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Penentuan nilai weighted sum dihitung dengan mengalikan weight dan input serta ditambahkan dengan nilai biases. Berdasarkan weighted sum, dapat diperoleh output dengan memasukkan nilai tersebut dalam fungsi aktivasi ReLu. ReLu akan mengubah nilai output dalam kisaran 0 hingga infinity. Perhitungan menggunakan persamaan 2.5. Berikut hasil perhitungan neuron pada hidden layer pertama.



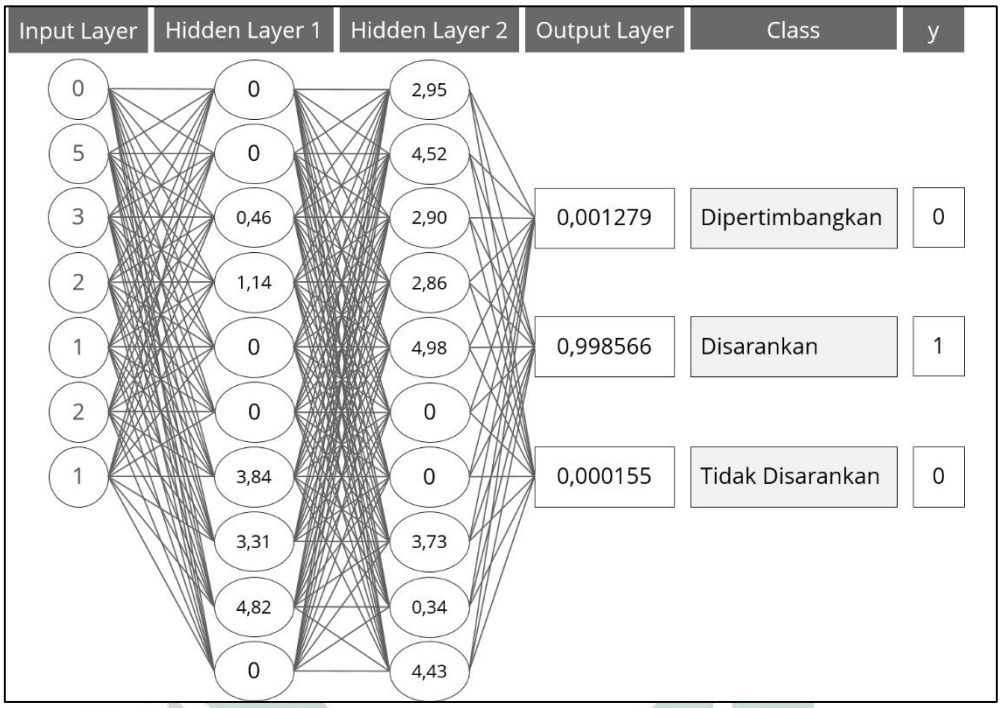
Gambar 4.2 Hasil Perhitungan Neuron Pada Hidden Layer

Neuron selanjutnya dihitung dengan cara yang sama, sedangakan penentuan neuron untuk output layer menggunakan menggunakan activation function *Softmax*. Softmax menghasilkan probabilitaas setiap class yang jika dijumlahkan harus sama dengan 1. Probabilitas tertinggi menunjukkan output class.

$$\begin{vmatrix} 0{,}26 & -0{,}67 & 0{,}84 & 0{,}8 & 0{,}6 & 1{,}86 & -0{,}39 \\ -1{,}04 & 0{,}61 & -1{,}73 & -0{,}23 & 0{,}88 & -2{,}36 & -0{,}59 \\ 0{,}27 & 0{,}15 & 1{,}83 & -0{,}95 & -1{,}21 & -2{,}55 & -0{,}09 \end{vmatrix} \begin{vmatrix} 2{,}29 \\ 4{,}52 \\ 2{,}90 \\ 2{,}86 \\ 4{,}98 \\ 0 \\ 0 \\ 3{,}73 \\ 0{,}34 \\ 4{,}43 \end{vmatrix} + \begin{vmatrix} 0{,}17 \\ -0{,}59 \\ 0{,}59 \end{vmatrix} = \begin{vmatrix} -0{,}89 \\ -2{,}47 \\ 0{,}46 \end{vmatrix}$$

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berdasarkan weighed sum, dapat diperoleh output menggunakan persamaan 2.4 yaitu persaman fungsi aktivasi softmax. Dalam persamaan 2,4, nilai probabilitas dihitung berdasarkan nilai eksponensial suatu weighted sum dibagi dengan jumlah nilai eksponensial seluruh weighted sum.



Gambar 4.3 Hasil Perhitungan Neuron Pada Ouput Layer

Perhitungan cross entropy dilakukan untuk mengetahui kinerja model ANN. Cross entropy untuk multiclass classification menghitung nilai loss secara terpisah untuk setiap class dan menjumlahkan seluruh nilai loss. Perhitungan menggunakan persamaan 2.16.

$$L = -\sum_{c=1}^{M} y_c \, log(p_c) = -(0 - 0{,}00062 + 0) = 0{,}00062$$

Nilai loss yang dihasilkan adalah 0,00062. Untuk mengurangi loss tersebut backpropagation dilakukan dengan memperbarui weight menggunakan mekanisme metode optimasi. Penelitian ini menggunakan metode optimasi *classical Adam* dan kombinasi Adam dan Nesterov Accelerated Gradient.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Perhitungan weight mengunakan Adam diawali dengan inisialisasi gradien. Fungsi gradient diperoleh dari nilai turunan loss function. Adam menggunakan nilai gradients g_t, β_1 dan β_2 untuk menghitung rata-rata bergerak. β_1 adalah exponential decay untuk moment pertama dengan nilai default 0.9. β_2 adalah exponential decay moment kedua dengan nilai default 0.999. Perhitungan ini menggunakan persamaan 2.6 sampai 2.9. Berdasarkan nilai didapatkan, dapat diperoleh permbaruan weight.

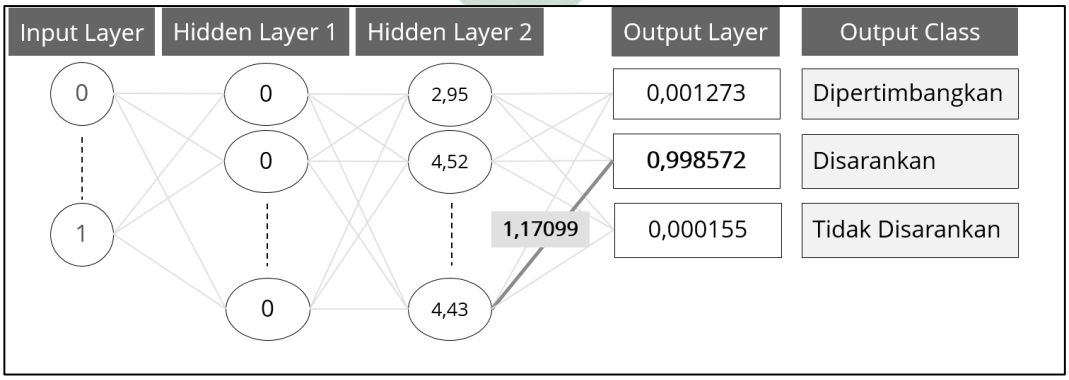
$$w_t = w_{t-1} - \eta \frac{m_t}{\sqrt{\hat{v}_t} + \varepsilon} = 1{,}17 - 0{,}001 \frac{-0{,}0015}{\sqrt{23 \times 10^{-7}} + 10^{-8}} = 1{,}17099$$

Weight dengan nilai awal 1,17 akan diperbarui oleh Adam menjadi 1,1709. Dalam pengembangan metode kombinasi Adam dan Nesterov, dilakuka modifikasi pada perhitungan momentum dengan persamaan 2.14.

$$w_{t+1} = w_t - \frac{\eta}{\sqrt{\hat{v}_t} + \epsilon} \left(\beta_1 \hat{m}_{t-1} + \frac{(1 + \beta_1) g_t}{1 - \beta_1^t} \right)$$

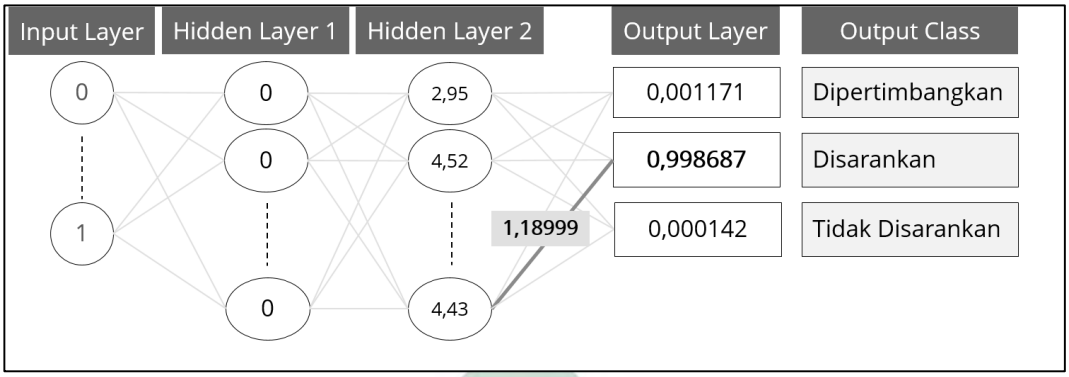
$$w_{t+1} = 1{,}17 - 0{,}001 \frac{-0{,}0285}{\sqrt{23 \times 10^{-7}} + 10^{-8}} = 1{,}18999$$

Metode kombinasi Adam dan Nesterov memperbarui weight menjadi 1,899. Nilai probabilitas akan berubah dan nilai loss kemungkinan akan menurun setelah pembaruan nilai weight dilakukan. Berikut nilai probabilitas yang dihasilkan Adam dan kombinasi Adam dan Nesterov.



Gambar 4.4 Hasil Pembaruan Weight Menggunakan Adam

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id



Gambar 4.5 Hasil Pembaruan Weight Menggunakan Adam dan Nesterov

Hasil perubahan weight menunjukkan bahwa kenaikan probabilitas tertinggi pada class Disarankan didapatkan oleh metode kombinasi Adam dengan Nesterov. Untuk mengetahui kinerja kedua metode tersebut, perhitungan nilai loss dilakukan.

$$Adam\ L = -\sum_{c=1}^{M} y_c\, log(p_c) = -(0 - 0{,}00062 + 0) = \mathbf{0{,}00062}$$

$$Adam + Nesterov\ L = -\sum_{c=1}^{M} y_c\, log(p_c) = -(0 - 0{,}00057 + 0) = \mathbf{0{,}00057}$$

Perhitungan loss menunjukan bahwa pembaruan nilai weight menggunakan metode kombinasi Adam dan Nesterov Momentum dapat menghasilkan loss lebih rendah daripada Adam dalam pemodelan prediksi kesiapan sekolah secara manual.

Untuk menguji kinerja kombinasi Adan dan Nesterov, penelitian melakukan pelatihan model dengan beberapa skenario jumlah neuron dan iterasi. Penelitian ini menggunakan Scikit-Learn dan Grid Search cv untuk memudahkan pengujian. Grid Search alat bantu untuk menemukan skenario parameters Artificial Neural Network dengan performance terbaik.

```
INIT classifier TO mlpClassifier
SET parameters, numberNeuron, maxIter, solver, nesterovMomentum
WHILE gridSearch is use
      CALL classifier, parameters
      SET crossValidation to 5, scoring to accuracy
      FIT featuresData, targetData
ENDWHILE
```


digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Skenario pengujian menggunakan jumlah neuron angka genap dimulai dari nilai 6 - 22, sedangkan skenario epochs menggunakan dua nilai yaitu 500 dan 1000. Pengujian menggunakan skenario pembagian data *5-fold cross validation*.

Variable feature dan variable target classes pada data hasil kesiapan sekolah dideskripsikan untuk proses training model atau classification data. Hasil pengujian ditunjukkan sebagai berikut.

```
Best Param : {'hidden_layer_sizes': (10, 10),
'max_iter': 1000,'nesterovs_momentum': True,
'solver': 'adam'}

Best Score : 0.96
```

Hasil pengujian menunjukkan bahwa akurasi terbaik diperoleh model ANN dengan jumlah neuron pada hidden layer adalah 10, jumlah iterasi adalah 1000 dan menggunakan metode kombinasi Adam dan Nesterov Momentum. Arsitektur ANN tersebut menghasilkan nilai akurasi sebesar 96%.

4.3.2 *Model Evaluation*

Evaluasi model menggunakan accuracy, time dan loss. Akurasi adalah rasio prediksi benar dengan jumlah seluruh data berdasarkan hasil confusion matrix. Hasil akurasi metode optimizations Adam dan Adam dengan Nesterov Momentum dibandingkan dalam Tabel 4.6.

Tabel 4.6 Hasil *Accuracy ANN* dengan *Adam* dan *Adam Nesterov*

No	***Architecture***	***Epochs***	***Accuracy***	
			Adam	***Adam + Nesterov***
1	7-6-6-1	500	0.8533	**0.8889**
2	7-6-6-1	1000	0.9244	**0.9467**
3	7-8-8-1	500	0.8978	**0.9111**
4	7-8-8-1	1000	**0.9333**	0.9289
5	7-10-10-1	500	0.9111	**0.9244**
6	7-10-10-1	1000	0.9511	**0.9600**

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.7 Hasil *Accuracy ANN* dengan *Adam* dan *Adam Nesterov* (lanjutan)

No	***Architecture***	***Epochs***	***Accuracy***	
			Adam	***Adam + Nesterov***
7	7-12-12-1	500	**0.9467**	0.9200
8	7-12-12-1	1000	0.9511	0.9511
9	7-14-14-1	500	**0.9556**	0.9467
10	7-14-14-1	1000	0.9511	0.9511
11	7-16-16-1	500	0.9289	**0.9467**
12	7-16-16-1	1000	0.9422	**0.9467**
13	7-18-18-1	500	0.9467	**0.9556**
14	7-18-18-1	1000	**0.9467**	0.9422
15	7-20-20-1	500	0.9467	0.9467
16	7-20-20-1	1000	0.9467	**0.9511**
17	7-22-22-1	500	0.9422	**0.9467**
18	7-22-22-1	1000	0.9556	0.9556

Berdasarkan hasil accuracy pada Tabel 4.6, tingkat accuracy tertinggi adalah sebesar 0.9600 didapatkan oleh metode kombinasi Adam dan Nesterov Accelereted Gradient dengan architecture model 10 neurons pada hidden layer dan 1000 epochs. Tingkat akurasi terendah sebesar 0.8533 diperoleh metode Adam dengan arsitektur 6 neurons pada hidden layer dan 500 epochs.

Hasil pengujian menunjukan bahwa nilai epoch memiliki pengaruh terhadap tingkat akurasi model. Semakin besar *epochs*, semakin tinggi tingkat *accuracy* yang akan didapatkan. Namun, pengembangan jumlah neurons pada hidden layers tidak memiliki pengaruh terhadap peningkatan nilai accuracy.

Evaluasi model juga dilakukan menggunakan ukuran waktu atau kecepatan. Waktu yang dimaksud adalah lamanya model dalam melakukan pelatihan data dan pengujian data. Tabel 4.8 menunjukkan hasil model mean time yaitu jumlah waktu rata-rata untuk seluruh data fold pada setiap skenario parameter.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tabel 4.8 Hasil *Modeling Time Adam* dan *Adam Nesterov*

No	***Architecture***	***Epochs***	***Time* (*second*)**	
			Adam	***Adam + Nesterov***
1	7-6-6-1	500	**1.3313**	1.3799
2	7-6-6-1	1000	**2.5469**	2.6157
3	7-8-8-1	500	1.3754	**1.3638**
4	7-8-8-1	1000	2.3376	**2.2251**
5	7-10-10-1	500	1.3656	**1.3719**
6	7-10-10-1	1000	2.3876	**1.9881**
7	7-12-12-1	500	1.4406	**1.4250**
8	7-12-12-1	1000	**2.0813**	2.1813
9	7-14-14-1	500	**1.4625**	1.4719
10	7-14-14-1	1000	1.9438	**1.8844**
11	7-16-16-1	500	1.5219	1.5219
12	7-16-16-1	1000	1.8844	**1.8313**
13	7-18-18-1	500	1.5313	**1.4875**
14	7-18-18-1	1000	1.6375	**1.6000**
15	7-20-20-1	500	**1.6356**	1.7449
16	7-20-20-1	1000	1.7657	**1.5938**
17	7-22-22-1	500	**1.5250**	1.5375
18	7-22-22-1	1000	1.6313	**1.6063**

Berdasarkan hasil pada 4.8, waktu tercepat dalam melakukan training 1,331 second diperoleh metode Adam dengan architecture model 6 neuron dan 500 epoch. Namun, architecture tersebut memiliki tingkat accuracy paling rendah yaitu sebesar 0.8533 dalam proses model evaluation sebelumnya. Sehingga, architecture tersebut dapat dikatakan belum optimal.

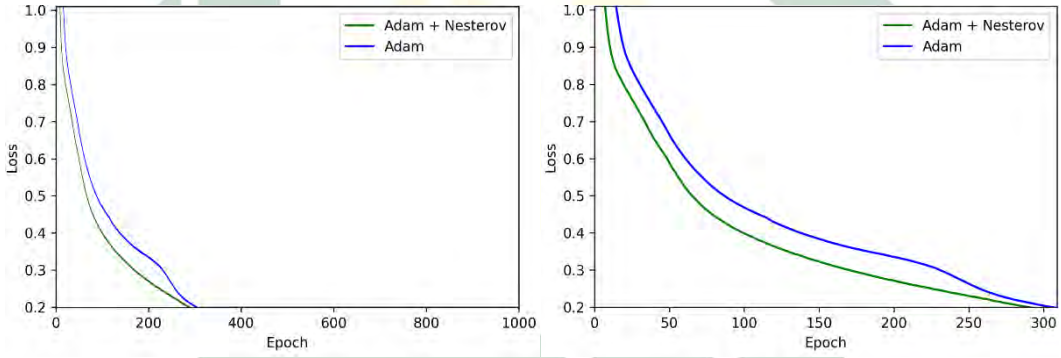
Sebelumnya, akurasi paling tinggi diperoleh kombinasi Adam dan Nesterov dengan architecture model 10 neurons pada hidden layer dan 1000 epochs. Dalam evaluasi waktu, model tersebut juga memiliki waktu yang relatif cepat dalam proses training yaitu 1,988 seconds. Sehingga, model tersebut dapat dikatakan optimal.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Evaluasi model selanjutnya menggunakan nilai loss function. Loss function digunakan untuk menghitung nilai *error* dengan mengukur perbedaan antara output prediksi dan output yang diinginkan. Loss functions yang digunakan dapat berupa categorical cross entropy untuk klasifikasi multiclass.

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 model Adam dan Adam dengan Nesterovs menggunakan skenario architecture model dengan 10 neuron pada *hidden layer* dan 1000 *epochs*. Namun, sebelum mencapai 1000 epochs, model telah menghasilkan minimum loss pada 300 epochs. Hasil loss function yang didapatkan pada setiap *epoch* digambakan dalam bentuk kurva untuk memudahkan dalam membandingkan kinerja metode.



Gambar 4.6 Kurva *Loss Function* Metode *Adam* dan *Adam Nesterov*

Kurva tersebut menunjukkan bahwa *loss function* paling rendah didapatkan oleh metode kombinasi Adam dan Nesterov Accelereted Gradient. Optimization ini memiliki *loss function* lebih rendah daripada Adam pada setiap *epochs*. Minimum loss juga lebih cepat didapatkan oleh Adam Nesterov dibandingkan dengan Adam.

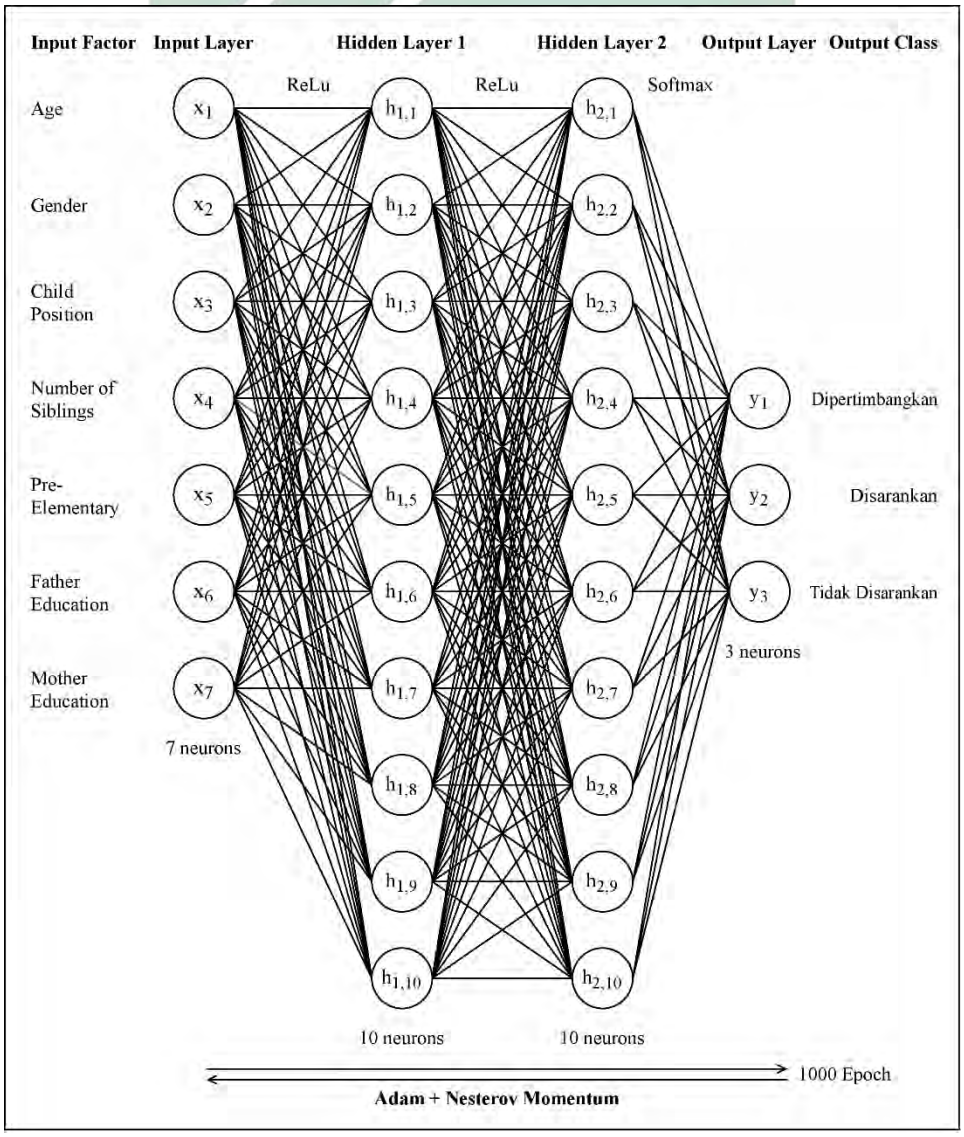
Tabel 4.9 Hasil *Loss Function ANN* dengan *Adam* dan *Adam Nesterov*

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**

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Berdasarkan hasil *loss function* pada Tabel 4.9, nilai *loss function* terendah adalah sebesar 0.04 didapatkan oleh metode Adam dengan Nesterov's Accelerated Gradient. Metode optimization ini memiliki mean *loss function* lebih baik daripada Adam yaitu senilai 0.062 untuk Adam Nesterov dan 0.068 untuk Adam.

Hasil evaluasi Adam dan Adam dengan Nesterov Momentum menggunakan nilai accuracy, time dan loss menunjukkan bahwa kinerja terbaik dilakukan oleh metode kombinasi Adam dan Nesterov Accelereted Gradient. Metode ini memiliki accuracy 96% dengan waktu training 1,988 seconds dan loss senilai 0.062 dengan architecture 10 neuron pada hidden layer dan 1000 epochs.

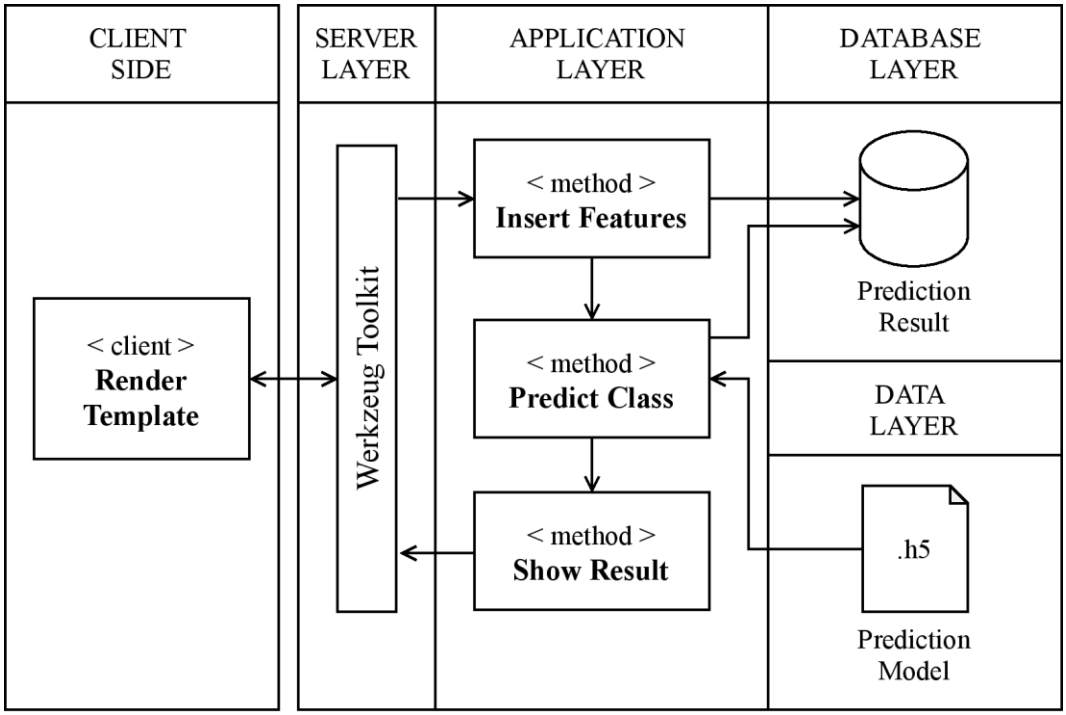


Gambar 4.7 Arsitektur Model *Artificial Neural Network* Terbaik

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4.4 *Deployement*

Deployment adalah proses penyaluran machine learning untuk dimanfatkan dalam kehidupan. Flask adalah framework yang digunakan dalam pengembangan sistem prediksi kesiapan sekolah berbasis web. Architecture pengembangan model prediksi berbasis *website* ditunjukkan pada Gambar 4.9.



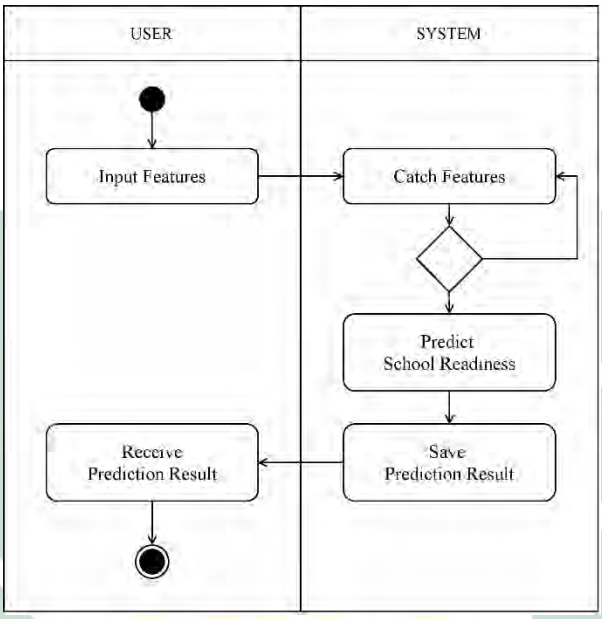
Gambar 4.8 Arsitektur Sistem Prediksi

Flask memiliki komponen *Jinja* dan *Werkzeug*. Jinja merupakan komponen untuk menampilan web, sedangkan werkzeug adalah komponen untuk web servers. *Server side* terdiri dari komponen *data*, *database* dan *application*. *Data layer* adalah komponen yang memuat model prediksi kesiapan sekolah, sedangkan *databse layer* berfungsi untuk menyimpan data hasil prediksi.

Application layer mendefinisikan fungsi dalam *flask application*. Sistem ini dapat menangkap masukan *features*, memuat model prediksi, melakukan prediksi berdasarkan *features*, menyimpan hasil prediksi dalam *database* dan menampilkan riwayat prediksi kepada pengguna. *Database layer* menggunakan *MySQL*. Aktifitas utama pengguna terdiri dari melakukan prediksi dan menampilkan *history*. Hal ini dapat digambarkan melalui *activity diagram*.

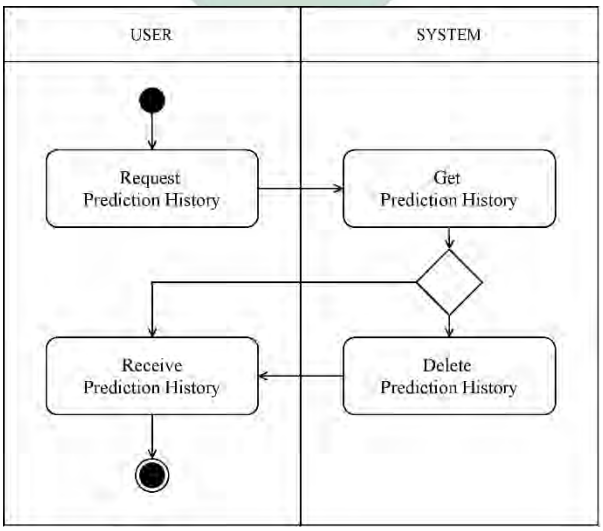
digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Activity diagram untuk proses prediksi. Prediksi adalah proses memasukkan data *features* yang terkait dengan data diri peserta didik. Data diri terdiri dari nama, jenis kelamin, tanggal lahir, jumlah saudara, pendidikan ayah, pendidikan ibu, kelas dan urutan anak. Data tersebut digunakan untuk memprediksi kesiapan sekolah.



Gambar 4.9 *Activity Diagram* Prediksi

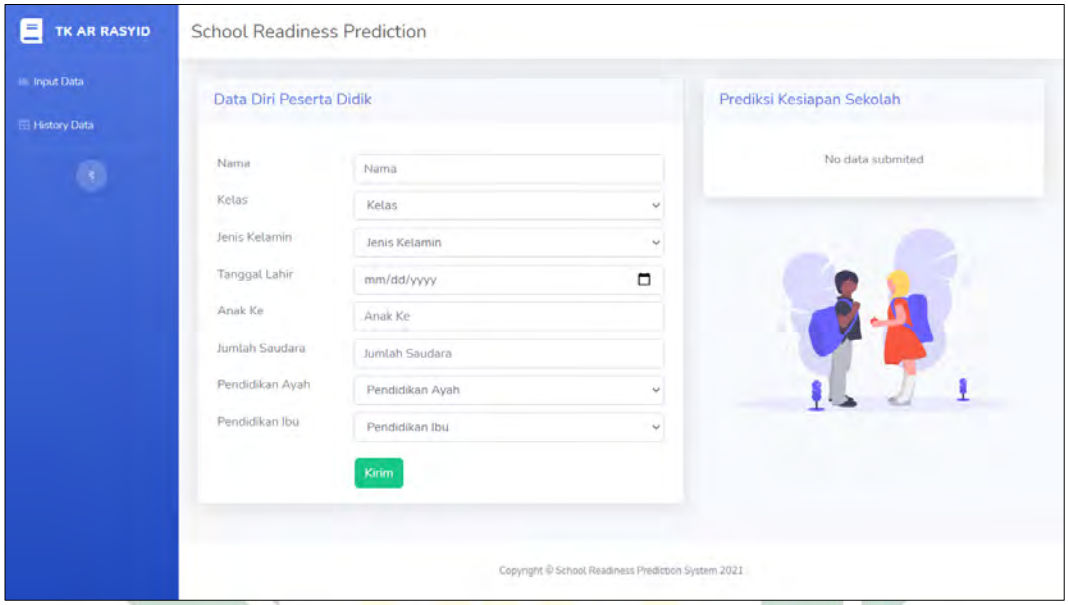
Activity diagram untuk menampilkan data history. History merupakan data riwayat prediksi yang terdiri dari *features* dan *prediction*. Data tersebut ditampilkan kepada pengguna, selanjutnya pengguna dapat melakukan penghapusan data.



Gambar 4.10 *Activity Diagram* History

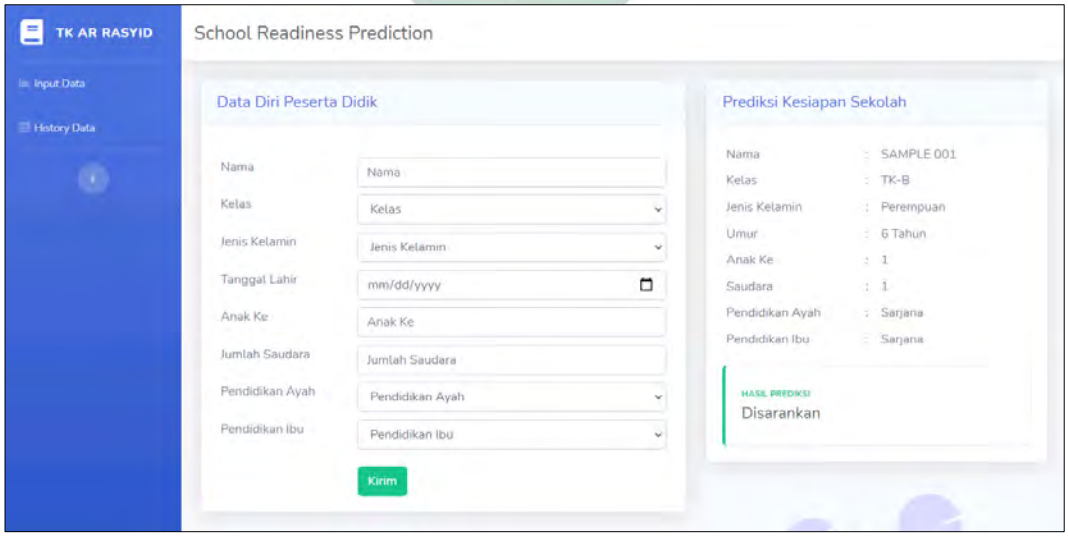
digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Hasil implementasi sistem ini berdasarkan *architecture* dan *activity diagram* dapat ditunjukkan melalui *user interface*. *User interface* adalah media penghubung antara sistem dengan pengguna. Sistem prediksi kesiapan sekolah meliputi tampilan menu *input* data dan *history* data. Berikut adalah tampilan dari sistem.



Gambar 4.11 *User Interface* Input Data

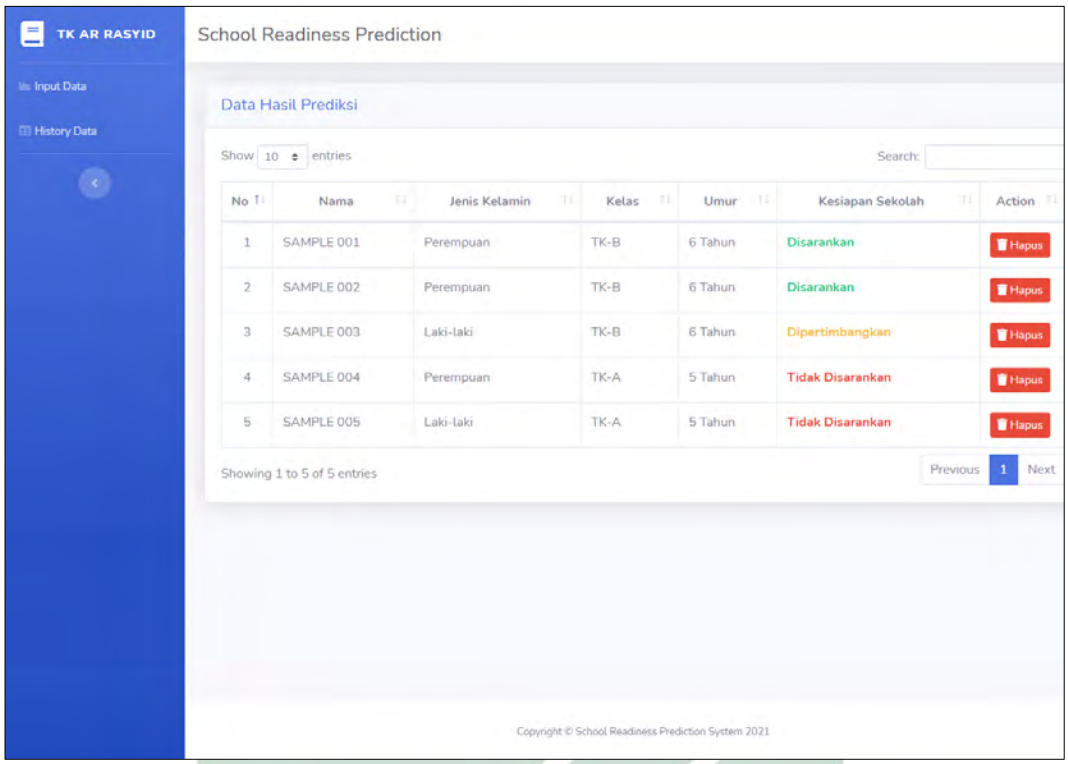
Gambar 4.12 adalah tampilan dari *form* input data. *Form* untuk memasukan *features* prediksi yang terdiri dari nama, kelas, jenis kelamin, tanggal lahir, anak ke, jumlah saudara, pendidikan ayah dan pendidikan ibu.



Gambar 4.12 *User Interface* Hasil Prediksi

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Gambar 4.13 adalah tampilan dari hasil prediksi. Hasil prediksi menampilka klasifikasi berupa disarankan, dipertimbangkan atau tidak disarankan untuk masuk sekolah dasar. Klasifikasi menampilkan warna untuk masing-masing *class*. Warna hijau untuk *clas* disarankan, warna kuning untuk dipertimbangkan dan warna merah untuk *class* tidak disarankan. Selain itu, terdapat data *features* prediksi.



Gambar 4.13 *User Interface* History Data

Gambar 4.14 terdapat tampilan dari *history* data. *History* data menampilkan riwayat hasil prediksi meliputi nama, jenis kelamin, kelas, umur dan hasil prediksi kesiapan sekolah. Data history digunakan untuk melakukan evaluasi. *User interface* menunjukkan bahwa setiap hasil prediksi menampilkan warna tertentu.

Action yang dapat dilakukan oleh pengguna dalam menu *history* data adalah mendapatkan data dari database dan melakukan penghapusan data. *User* tidak dapat mengubah data *features* dan hasil prediksi kesiapan sekolah. Implementasi sistem prediksi selanjutnya akan melalui proses pengujian untuk mengetahui nilai akurasi dengan membandingkan hasil prediksi dengan *NST*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Pengujian berfungsi untuk membandingkan hasil prediksi berbasis *Artificial Neural Network* dengan hasil tes kesiapan sekolah berbasis *NST*. Jumlah data yang digunakan sebanya 225 data yang terdiri dari 3 *class label*. Tabel 4.10 menunjukkan hasil pengujian sistem menggunakan *ANN* dengan metode optimasi *Adam*.

Tabel 4.10 Hasil Pengujian Sistem menggunakan *Adam*

Class Label	***Total***	***True***	***False***	***Accuracy***
Disarankan	133	131	2	98.85%
Dipertimbangkan	27	25	2	92.59%
Tidak Disarankan	65	63	2	96.92%
Total	**225**	**219**	**6**	**97.33%**

Hasil pengujian sistem menggunakan algoritma *Adam* menunjukkan angka *accuracy* sebesar 97.33%. Jumlah data dengan prediksi benar adalah 219 dan data dengan prediksi salah adalah 6 data. Berikut adalah hasil pengujian sistem dengan metode kombinasi *Adam* dan *Nesterov* momentum.

Tabel 4.11 Hasil Pengujian Sistem menggunakan *Adam* dan *Nesterov*

Class Label	***Total***	***True***	***False***	***Accuracy***
Disarankan	133	131	2	98.23%
Dipertimbangkan	27	26	1	86.95%
Tidak Disarankan	65	64	1	96.92%
Total	**225**	**221**	**4**	**98.22%**

Pengujian sistem menggunakan metode kombinasi *Adam* dengan *Nesterovs* momentum menghasilkan nilai *accuracy* sebesar 98.22%. Sistem prediksi dengan algoritma kombinasi Adam dengan Nesterov mampu memprediksi 221 data dengan benar dan sejumlah 7 data dengan prediksi tidak benar atau gagal.

Berdasarkan pengujian mengunakaan algoritma *Adam* dan *Adam Nesterov*, hasil menunjukkan bahwa algoritma kombinasi Adam dengan Nesterov momentum memiliki nilai accuracy lebih baik dibandingkan dengan Adam dalam memprediksi kesiapan sekolah berdasarkan data hasil tes kesiapan sekolah berbasis *NST*.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Adam dan kombinasi Adam dengan Nesterov masih menghasilkan prediksi yang salah untuk beberapa jumlah data. Incorrect prediction dapat disebabkan oleh machine learning model atau datates yang digunakan. Tabel 4.12 menunjukkan data dengan incorrect prediction dalam metode Adam.

Tabel 4.12 *Incorrect Prediction* Metode *Adam*

Gender	***Age***	***Position***	***Siblings***	***Preschool***	***Father Edu***	***Mother Edu***	***Class***	***Predict***
P	6	3	2	Tidak	SMP	SMP	Dipertimbangkan	Tidak Disarankan
P	6	1	1	Tidak	S1	SMA	Dipertimbangkan	Tidak Disarankan
L	5	2	1	Pernah	S1	S1	Tidak Disarankan	Dipertimbangkan
L	5	2	1	Pernah	S1	SMA	Tidak Disarankan	Disarankan
L	5	2	1	Pernah	S1	S1	Disarankan	Dipertimbangkan
L	5	2	1	Pernah	S1	S1	Disarankan	Dipertimbangkan

Berdasarkan data tersebut, dapat diketahui bahwa incorect prediction terjadi akibat adanya inconsistent data. Data menunjukkan nilai variable feature yang sama mempunyai variable target *class* yang berbeda.

Tabel 4.13 *Incorrect Prediction* Metode *Adam Nesterov*

Gender	***Age***	***Position***	***Siblings***	***Preschool***	***Father Edu***	***Mother Edu***	***Class***	***Predict***
L	5	2	1	Pernah	S1	S1	Tidak Disarankan	Dipertimbangkan
L	5	2	1	Pernah	SMA	SMA	Dipertimbangkan	Disarankan
L	5	2	1	Pernah	S1	S1	Disarankan	Dipertimbangkan
L	5	2	1	Pernah	S1	S1	Disarankan	Dipertimbangkan

Incorrect predictions metode kombinasi juga disebabkan inconsistent data. Metode kombinasi Adam dan Nesterov lebih baik dalam handling inconsistent data, sehingga nilai incorrect prediction lebih rendah daripada Adam.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

4.5 Diskusi

Penelitian ini bertujuan untuk mengetahui kinerja metode kombinasi *Adam* dan *Nesterov Momentum* dibandingkan dengan *Adam* dalam memprediksi kesiapan sekolah. Hasil penelitian menunjukkan bahwa algoritma kombinasi *Adam* dengan *Nesterov Momentum* memiliki kinerja lebih baik daripada *Adam Optimizer* dalam memprediksi kesiapan sekolah dengan akurasi 96% dalam waktu 1,9881 seconds.

Nesterov Momentum digunakan untuk memperbaiki *momentum* pada *Adam*. *Nesterov* menggunakan *current parameter* untuk mengetahui nilai *future loss*. Jika pembaruaan *parameters* menyebabkan *bad loss*, maka *gradien* akan mengarahkan pembaruan kembali ke parameter sebelumnya. Hal ini membantu *Nesterov* untuk menghindari *oscillation* atau *repetitive variation* yang tidak dapat dilakukan oleh *vanilla momentum* (Goodfellow, Bengio, & Courville, 2016).

Dengan demikian, *Nesterov* dapat mencapai nilai minimum loss lebih cepat dibandingkan dengan vanilla momentum. Hasil penelitiaan ini menunjukkan bahwa algoritma kombinasi *Adam* dengan *Nesterov Momentum* memiliki loss value lebih rendah daripaada *Adam Optimizer* yang menggunakan vanilla momentum dengan nilai loss function kombinasi *Adam* dan *Nesterov* 0.062 dan *Adam* 0.068.

Penelitian ini juga bertujuan untuk mengetahui akurasi *Adam* dan algoritma kombinasi *Adam* dan *Nesterov Momentum* jika dibandingkan dengan hasil tes *NST*. Hasil penelitian ini menunjukkan bahwa prediksi kesiapan sekolah menggunakan *Adam* memiliki akurasi 97.33% terhadap hasil *NST*, sedangkan metode kombinasi *Adam* dan *Nesterov Momentum* mempunyai akurasi lebih baik dibandingkan *Adam* dengan akurasi sebesar 98.22% terhadap hasil tes kesiapan sekolah berbasis *NST*.

Hasil penelitian menunjukkan bahwa metode *Nesterovs Momentum* mampu meningkatkan kinerjaa *Adam* dalam memprediksi kesiapaan sekolah. Hal tersebut sejalan dengan hasil penelitian Dozat yang melakukan komparasi terhadap enam optimization algorithm yaitu *SGD*, *Momentum*, *NAG*, *RMSProp*, *Adam* dan *Nadam*. Hasil menunjukkan bahwa metode kombinasi *Adam* dengan *Nesterov Momentum* mengungguli algoritma lainnya termasuk algoritma induknya, *Adam*. Algoritma ini paling unggul dalam menurunkan *traning* dan *validation loss* (Dozat, 2016).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

BAB V
PENUTUP

5.1 Kesimpulan

Berdasarkan penelitian yang telah dilaksanakan, didapatkan hasil penelitian terkait algoritma kombinasi *Adam* dan *Nesterov Momentum* dalam pengembangan model prediksi kesiapan sekolah yang dapat disimpulkan sebagai berikut.

1. Kombinasi *Adam* dan *Nesterov Momentum* (*Nadam*) memiliki kinerja lebih baik dibandingkan Adam dalam memprediksi kesiapan sekolah. Kombinasi metode tersebut menghasilkan *accuracy* sebesar 96% dan *loss* senilai 0.06 dalam 1,98 *seconds* dengan 10 *neuron* pada 2 *hidden layer* dan 1000 *epochs*.
2. *Adam* dengan *Nesterov Momentum* juga menghasilkan ketepatan lebih baik dibandingkan *Adam* dalam memprediksi kesiapan sekolah sesuai hasil *NST* dengan nilai ketepatan sebesar 98.22%.

5.2 Saran

Penelitian ini yang terkait dengan prediksi kesiapan sekolah maupun metode optimization masih memerluka adanya pengembangan. Penelitian berikutnya dapat dilakukan menurut saran berikut.

1. Prediksi kesiapan sekolah dapat ditingkatkan dengan menambakan variable *feature* lainnya, misalnya family size, parents occupation, neighborhood dan class peers (Suleiman et al., 2017).
2. Metode optimization dapat menggunakan *Accelerated Adam* yang mampu meningkatkan kinerja *Adam* dan *Nadam* (Tato & Nkambou, 2018).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

DAFTAR PUSTAKA

Al-Jauziyah, S. I. Q. (1990). *'Aun Al-Ma'bud Syarah Sunan Abu Daud*. Beirut: Darul Kutb Al-Ilmiyah.

Alasadi, S., & Bhaya, W. (2017). Review of Data Preprocessing Techniques in Data Mining. *Journal of Engineering and Applied Sciences*, *12*(16), 4102–4107.

Blank, J., & Deb, A. (2020). Pymoo: Multi Objective Optimization in Python. *IEEE2*, *8*, 89497–89509. https://doi.org/10.1109/ACCESS.2020.2990567

Bock, S., Goppold, J., & Wei, M. (2018). An Improvement of The Convergence Proof of The Adam Optimizer. In *OTH Clusterkonferenz* (pp. 1–3). Weiden: Cornell University.

Calders. (2012). Cost-Sensitive Classification Problem. In *Workshop on Teching Machine Learning*. Edinburgh: International Conference on Machine Learning.

Dozat, T. (2016). Incorporating Nesterov Momentum Into Adam. *ICLR*.

El-Khatib, M. J. (2019). Glass Classification Using Artificial Neural Network. *International Journal of Academic Pedagogical Research*, *3*(2).

Fauziyah, N. (2019). School Readiness of Preschool Children Entering Elementary School Based on Nijmeegse Schoolbekwaamheids Test: Case Study at Early Childhood Education Al Husna Pengging Boyolali In 2017. *Early Childhood Research Journal*, *2*(2), 29–35.

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), 17–32.

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. Cambridge: MIT Press.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Gorunescu. (2011). *Data Mining Concept, Models and Techniques* (12th Editi). Craiova: Springer.

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).

Heidari, A. A., Faris, H., Aljarah, I., & Mirjalili, S. (2018). An Efficient Hybrid Multilayer Perceptron Neural Network with Grasshopper Optimization. *Soft Computing*. https://doi.org/10.1007/s00500-018-3424-2

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), 40–46.

Huda, N. S., Mubarok, M. S., & Adiwijaya. (2019). A Multi-label Classification on Topics of Quranic Verses (English Translation) Using Backpropogation 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. https://doi.org/10.1109/CITSM47753.2019.8965346

Khan, S. A., & Rana, Z. A. (2019). Evaluating Peformance of Software Defect Prediction Models Using Area Under Precision-Recall Curve (AUC-PR). In *International Conference on Advancements in Computational Sciences* (pp. 1–6). Lahore: IEEE.

Kingman, D. P., & Ba, J. (2015). Adam: A Method for Stochastic Optimization. In *3rd International Conference for Learning Representations*. San Diego.

Li, Z., Zhang, Q., & Zhao, X. (2017). Performance Analysis of K-Nearest Neighbor, Support Vector Machine and Artificial Neural Network Classifier for Driver

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Drowsiness Detection with Different Road Geometries. *International Journal of Distributed Sensor Networks*, *1*(9). https://doi.org/10.1177/1550147717733391

Mendikbud. (2018). *Peraturan Menteri Pendidikan Dan Kebudayaan Republik Indonesia Nomor 14 Tahun 2018 Tentang Penerimaan Peserta Didik Baru Pada Taman Kanak-Kanak, Sekolah Dasar, Sekolah Menengah Pertama, Sekolah Menengah Atas, Sekolah Menengah Kejuruan, Atau Bentuk Lainnya*.

Monks, Rost, & Coffie. (1978). *Manual Nijmeegse Schoolbekwaamheids Test*.

Mufid, Rasyid, Rochimansyah, & Rokhim. (2019). Design an MVC Model Using Python for Flask Framework Development. *International Elctronics Symposium*, 214–219.

Nikam, S. (2015). A Comparative Study of Classification Techniques in Data Mining Algorithms. *Oriental Joural of Computer Science & Technology*, *8*(1), 13–19.

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.

Pradhan, N., Dhaka, V., & Kulhari, S. (2020). Experimental and Comparison Based Study on Diabetes Prediction Using Artificial Neural Network. *Recent Advances in Computer Science and Communications*, *13*(6), 1173–1179. https://doi.org/10.2174/2213275912666190801112119

Sabol, T., & Pianta, R. (2017). The State of Young Children in The United States: School Readiness. In *The Wiley Handbook of Early Childhood Development Programs, Practices and Policies* (First Edit). New Jersey: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118937334

Suleiman, I., Arslan, M., & Ridley, M. (2017). Prediction Model for School Readiness. *International Journal of Scientific & Engineering Research*, *8*(3).

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id

Tato, A., & Nkambou, R. (2018). Improving Adam Optimizer. In A. Rush (Ed.), *6th International Conference on Learning Representations*. San Diego: International Conference on Learning Representations.

Wang, Y., Zhou, P., & Zhong, W. (2018). An Optimization Strategy Based on Hybrid Algorithm of Adam and SGD. In *2rd International Conference on Electronic Information Technology and Computer Engineering*. Shanghai. https://doi.org/10.1051/matecconf/201823203007

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), 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), 38–44.

Zaheer, R., & Shaziya, H. (2019). A Study of The Optimimzation Algorithms in Deep Learning. In *3rd International Conference on Inventive Systems and Control* (pp. 536–539). Coimbatore: IEEE.

digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id digilib.uinsby.ac.id