

**RANCANG BANGUN APLIKASI ANALISIS PERFORMA
ISOLATION FOREST BERBASIS KONSEP *COMPUTER ASSISTED
PROBLEM BASED LEARNING (CAPBL)***



SKRIPSI

**Disusun Sebagai Salah Satu Syarat
untuk Memperoleh Gelar Sarjana Komputer
pada Departemen Ilmu Komputer/ Informatika**

**Disusun oleh:
Rina Dwi Hastuti
24010315130116**

**DEPARTEMEN ILMU KOMPUTER/ INFORMATIKA
FAKULTAS SAINS DAN MATEMATIKA
UNIVERSITAS DIPONEGORO**

2019

HALAMAN PERNYATAAN KEASLIAN SKRIPSI

Saya yang bertanda tangan di bawah ini :

Nama : Rina Dwi Hastuti

NIM : 24010315130116

Judul : Rancang Bangun Aplikasi Analisis Performa *Isolation Forest* Berbasis Konsep *Computer Assisted Problem Based Learning (CAPBL)*

Dengan ini saya menyatakan bahwa dalam skripsi ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan di dalam daftar pustaka.

Semarang, 20 Agustus 2019



Rina Dwi Hastuti
24010315130116

HALAMAN PENGESAHAN

Judul : Rancang Bangun Aplikasi Analisis Performa *Isolation Forest* Berbasis Konsep
Computer Assisted Problem Based Learning (CAPBL).

Nama : Rina Dwi Hastuti

NIM : 24010315130116

Telah diujikan pada sidang skripsi pada tanggal 8 Agustus 2019 dan dinyatakan lulus pada tanggal 8 Agustus 2019.

Semarang, 20 Agustus 2019

Mengetahui,

Ketua Departemen Ilmu Komputer/ Informatika

FSM Undip



Dr. Reno Kusumaningrum, S.Si., M.Kom
NIP. 198104202005012001

Panitia Penguji Skripsi

Ketua,

Beta Noranita, S.Si., M.Kom
NIP.197308291998022001

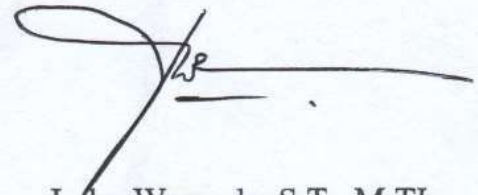
HALAMAN PENGESAHAN

Judul : Rancang Bangun Aplikasi Analisis Performa *Isolation Forest* Berbasis Konsep
Computer Assisted Problem Based Learning (CAPBL)
Nama : Rina Dwi Hastuti
NIM : 24010315130116

Telah diujikan pada sidang skripsi pada tanggal 8 Agustus 2019 dan dinyatakan lulus pada tanggal 8 Agustus 2019.

Semarang, 20 Agustus 2019

Pembimbing,



Indra Waspada, S.T., M.TI
NIP. 197902122008121002

KATA PENGANTAR

Puji syukur penulis panjatkan kehadiran Tuhan Yang Maha Esa yang telah melimpahkan segala rahmat dan hidayah-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul “Rancang Bangun Aplikasi Analisis Performa *Isolation Forest* Berbasis Konsep *Computer Assisted Problem Based Learning (CAPBL)* ”.

Dalam penyusunan laporan skripsi ini penulis mendapat banyak bimbingan, bantuan, dan dukungan dari berbagai pihak. Oleh karena itu, penulis menyampaikan terima kasih kepada:

1. Ibu Dr. Retno Kusumaningrum, S.Si., M.Kom, selaku Ketua Departemen Ilmu Komputer/ Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.
2. Bapak Panji Wisnu Wirawan, S.T., M.T., selaku Koordinator Skripsi Departemen Ilmu Komputer/ Informatika Fakultas Sains dan Matematika Universitas Diponegoro, Semarang.
3. Bapak Indra Waspada, S.T., M.TI., selaku dosen pembimbing skripsi yang telah membantu dalam membimbing dan mengarahkan penulis hingga selesainya skripsi ini.
4. Orang tua, keluarga, teman dekat dan sahabat yang telah mendukung, membantu, serta memberikan semangat kepada penulis dalam menyelesaikan skripsi ini.
5. Semua pihak yang telah membantu kelancaran dalam penyusunan skripsi, yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa dokumen skripsi ini masih jauh dari sempurna. Oleh karena itu, saran dan kritik yang membangun sangat penulis harapkan. Akhir kata, semoga skripsi ini dapat bermanfaat bagi semua pihak.

Semarang, 8 Agustus 2019

Rina Dwi Hastuti

HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS

Sebagai civitas akademik Universitas Diponegoro, saya yang bertanda tangan dibawah ini:

Nama : Rina Dwi Hastuti

NIM : 24010315130116

Program Studi : Informatika

Departemen : Ilmu Komputer/ Informatika

Fakultas : Sains dan Matematika

Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan **Hak Bebas Royalti Noneksklusif (*Non-exclusive Royalty Free Right*)** kepada Universitas Diponegoro atas karya ilmiah saya yang berjudul

Rancang Bangun Aplikasi Analisis Performa *Isolation Forest* Berbasis Konsep *Computer Assisted Problem Based Learning* (CAPBL)

beserta perangkat yang ada (jika diperlukan). Dengan Hak Bebas Royalti Non-eksklusif ini Universitas Diponegoro berhak menyimpan, mengalihmedia/ formatkan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan mempublikasikan skripsi saya tanpa meminta izin dari saya selama tetap mencantumkan nama saya sebagai penulis/ pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Semarang, 20 Agustus 2019

Yang menyatakan



Rina Dwi Hastuti

24010315130116

ABSTRAK

Deteksi anomali penting diterapkan pada berbagai bidang aplikasi untuk mengetahui kesalahan pada sebuah sistem. Dengan melakukan deteksi anomali maka dapat meminimalisir terjadinya kerugian pada kesalahan sistem tersebut. Deteksi anomali dapat dilakukan menggunakan beberapa algoritma. Salah satu algoritma tersebut adalah algoritma *Isolation Forest*. Algoritma *Isolation Forest* merupakan algoritma yang efisien dan efektif dalam menangani deteksi anomali. *Isolation Forest* memiliki performa yang lebih baik dari algoritma lain dalam hal waktu eksekusi, terutama dalam dataset yang besar. Meskipun algoritma *Isolation Forest* memiliki banyak kelebihan, namun masih sedikit *tools* yang menyediakan algoritma tersebut. Salah satu *tools* yang menyediakan algoritma *Isolation Forest* yaitu *Scikit Learn*. Akan tetapi, penggunaan *tools* tertentu seperti *Scikit Learn* membutuhkan waktu dan pengalaman yang cukup untuk mampu menguasai fitur-fitur pada *tools* tersebut. Dengan demikian, pada penelitian ini penulis mengembangkan sebuah aplikasi berbasis *web* yang digunakan untuk membantu pengguna dalam mempelajari serta mengamati performa dari *Isolation Forest* dengan disertai contoh dataset, pengaturan parameter *Isolation Forest*, visualisasi dan evaluasi. Aplikasi ini dikembangkan menggunakan model proses *waterfall* dan dikemas dengan CAPBL. Dalam menentukan fitur pada aplikasi dilakukan analisis fitur berdasarkan konsep CAPBL dengan proses bisnis dari metodologi CRISP-DM. Hasil pengujian dari aplikasi ini menyatakan bahwa aplikasi ini layak digunakan sebagai solusi baru untuk memfasilitasi pengguna yang ingin belajar serta mengamati performa *Isolation Forest*.

Kata Kunci : *Isolation Forest*, Deteksi Anomali, *Waterfall*, *Data Mining*, CAPBL, CRISP-DM

ABSTRACT

Anomaly detection was important applied in various fields of application to determine errors in the system. By detecting it can minimize losses on the system.. Anomaly detection can be done using several algorithms. One of these algorithm is the Isolation Forest algorithm. Isolation Forest Algorithm is a new algorithm that is efficient and effective in handling anomaly detection. Isolation Forests have better performance than other algorithm in terms of run time, especially in large datasets. Although the Isolation Forest algorithm has many advantages, there are still few tools that provide this algorithm. One tool that provides the Forest Isolation algorithm is Scikit Learn. However, using special tools such as Scikit Learning requires sufficient time and experience to be able to manage the features of the tool. Thus, in this study the authors developed a simple web-based application that is used to assist users in learning and managing the Isolation Forest by using a sample dataset, managing the Isolation Forest parameters, visualization and evaluation. This application was developed using the waterfall process model and packaged with CAPBL. In determining the features of the application, an analysis of features is based on the CAPBL concept with the business process of CRISP-DM approval. The test results of this application state that this application is fit to be used as a new solution to facilitate users who want to learn and support the Isolation Forest algorithm.

Keyword : *Isolation Forest, Anomaly Detection, Waterfall, Data Mining, CAPBL, CRISP-DM*

DAFTAR ISI

HALAMAN PERNYATAAN KEASLIAN SKRIPSI.....	ii
HALAMAN PENGESAHAN	iii
HALAMAN PENGESAHAN	iv
KATA PENGANTAR.....	v
HALAMAN PERNYATAAN PERSETUJUAN PUBLIKASI SKRIPSI UNTUK KEPENTINGAN AKADEMIS	vi
ABSTRAK	vii
ABSTRACT	viii
DAFTAR ISI	ix
DAFTAR TABEL	xii
DAFTAR GAMBAR.....	xiv
DAFTAR ISTILAH.....	xv
BAB I PENDAHULUAN	1
1.1. Latar Belakang.....	1
1.2. Rumusan Masalah.....	2
1.3. Tujuan dan Manfaat	2
1.4. Ruang Lingkup	3
1.5. Sistematika Penulisan	3
BAB II LANDASAN TEORI.....	5
2.1. Tinjauan Pustaka.....	5
2.2. Dasar Teori	7
2.2.1. Deteksi Anomali	7
2.2.2. Data Mining.....	8
2.2.3. Metodologi CRISP-DM.....	8
2.2.4. <i>Isolation Forest</i>	10
2.2.5. Deteksi Anomali menggunakan <i>Isolation Forest</i>	12
2.2.6. <i>Problem Based Learning (PBL)</i>	14
2.2.7. <i>Library Isolation Forest</i> pada <i>Scikit Learn</i>	15
2.2.8. <i>Principal Component Analysis (PCA)</i>	16
2.2.9. Metrik Evaluasi Performa Model Data Tidak Seimbang	20
2.2.10. <i>Waterfall</i>	21

2.2.11.	<i>Data and Process Modeling</i>	22
2.2.12.	Pengujian <i>Black Box</i>	22
BAB III	METODOLOGI PENELITIAN	23
3.1.	Studi Literatur	23
3.2.	Pengembangan Aplikasi	24
3.2.1.	Analisis CAPBL	24
3.2.2.	<i>Requirements Analysis and Definition</i>	25
3.2.1.1.	Analisis Tahap CRISP-DM	25
3.2.1.2.	Deskripsi Perangkat Lunak	27
3.2.1.3.	Karakteristik Pengguna	27
3.2.1.4.	Kebutuhan Fungsional	28
3.2.1.5.	Kebutuhan Non-Fungsional	28
3.2.3.	<i>System and Software Design</i>	28
3.2.2.1.	Pemodelan Data	28
3.2.2.2.	Pemodelan Fungsional	28
3.2.2.3.	Perancangan Antarmuka	28
3.2.4.	<i>Implementation and Unit Testing</i>	29
3.2.3.1.	Lingkungan Implementasi	29
3.2.3.2.	Implementasi Antarmuka	29
3.2.3.3.	Pengujian Perangkat Lunak	29
3.3.	Analisis Hasil	29
BAB IV	HASIL DAN PEMBAHASAN	30
4.1.	Pengembangan Aplikasi	30
4.1.1.	Hasil Analisis Persyaratan CAPBL	30
4.1.2.	<i>Requirement Analysis and Definition</i>	31
4.1.2.1.	Hasil Analisis Tahap CRISP-DM	31
4.1.2.2.	Deskripsi Perangkat Lunak	33
4.1.2.3.	Karakteristik Pengguna	34
4.1.2.4.	Kebutuhan Fungsional	34
4.1.2.5.	Kebutuhan Non-Fungsional	35
4.1.3.	<i>System and Software Design</i>	35
4.1.3.1.	Pemodelan Data	36
4.1.3.2.	Pemodelan Fungsional	36

4.1.3.3.	Perancangan Antarmuka	39
4.1.4.	Implementation and Unit Testing	43
4.1.4.1.	Lingkungan Implementasi.....	43
4.1.4.2.	Implementasi Antarmuka	44
4.1.4.3.	Rencana Pengujian	50
4.1.4.4.	Pelaksanaan Pengujian	51
4.1.4.5.	<i>Defect</i> dan Status Perbaikan.....	51
4.1.4.6.	Evaluasi Pengujian	51
4.2.	Analisa Hasil.....	51
BAB V	PENUTUP	68
5.1.	Kesimpulan	68
5.2.	Saran	68
DAFTAR	PUSTAKA.....	69
LAMPIRAN -	LAMPIRAN	71

DAFTAR TABEL

Tabel 2. 1 Perkembangan Penelitian Deteksi Anomali Menggunakan Algoritma Isolation Forest	5
Tabel 2. 2 Perkembangan Penelitian Pengembangan Aplikasi Menggunakan Konsep CAPBL	6
Tabel 2. 3 Tabel Perbandingan AUC Deteksi Anomali (Liu, Zhou, & Ting, 2008).	12
Tabel 2. 4 Tabel Waktu Pemrosesan Deteksi Anomali (Liu, Zhou, & Ting, 2008).....	12
Tabel 2. 5 Parameter Library Isolation Forest pada Tools Scikit Learn (scikit-learn developers, n.d.)	15
Tabel 2. 6 Confusion Matrix	20
Tabel 2. 7 Metrik <i>Threshold</i>	20
Tabel 4. 1 Hasil Analisis Persyaratan CAPBL.....	31
Tabel 4. 2 Fitur Hasil Identifikasi Persyaratan CRISP-DM	33
Tabel 4. 3 Tabel Daftar Aktor	34
Tabel 4. 4 Kebutuhan Fungsional Aplikasi Analisis Performa Isolation Forest Berbasis Konsep CAPBL	34
Tabel 4. 5 Kebutuhan Non-Fungsional Aplikasi Analisis Performa Isolation Forest Berbasis Konsep CAPBL	35
Tabel 4. 6 Hasil Perbandingan Parameter Number of Tree pada Percobaan Isolation Forest menggunakan Dataset Diabetes dengan Metode fit dan predict.....	57
Tabel 4. 7 Hasil Perbandingan Parameter Sample Size pada Percobaan Isolation Forest menggunakan Dataset Diabetes dengan Metode fit dan predict.....	58
Tabel 4. 8 Hasil Perbandingan Parameter Contamination pada Percobaan Isolation Forest menggunakan Dataset Diabetes dengan Metode fit dan predict.....	58
Tabel 4. 9 Hasil Perbandingan Parameter Number of Tree pada Percobaan Isolation Forest menggunakan Dataset Cardio dengan Metode fit dan predict.....	62
Tabel 4. 10 Hasil Perbandingan Parameter Sample Size pada Percobaan Isolation Forest menggunakan Dataset Diabetes dengan Metode fit dan predict.....	62
Tabel 4. 11 Hasil Perbandingan Parameter Contamination pada Percobaan Isolation Forest menggunakan Dataset Diabetes dengan Metode fit dan predict.....	63
Tabel 4. 12 Hasil Perbandingan Parameter Number of Tree pada Percobaan Isolation Forest menggunakan Dataset Satellite dengan Metode fit dan predict.	66

Tabel 4. 13 Hasil Perbandingan Parameter Sample Size pada Percobaan Isolation Forest menggunakan Dataset Diabetes dengan Metode fit dan predict.....	66
Tabel 4. 14 Hasil Perbandingan Parameter Contamination pada Percobaan Isolation Forest menggunakan Dataset Diabetes dengan Metode fit dan predict.....	66

DAFTAR GAMBAR

Gambar 2. 1 Metodologi CRISP-DM (Larose & Larose, 2014)	9
Gambar 2. 2 Persentase varian dari setiap PCs	18
Gambar 2. 3 Model Proses Waterfall (Sommerville, 2010)	21
Gambar 3. 1 Metodologi Penelitian.....	23
Gambar 4. 1 PCA Diabetes.....	32
Gambar 4. 2 DCD Aplikasi Analisis Performa Isolation Forest Berbasis Konsep CAPBL	36
Gambar 4. 3 DFD Level 1 Aplikasi Analisis Performa Isolation Forest Berbasis Konsep CAPBL	37
Gambar 4. 4 DFD Level 2 dari Proses Informasi Data	38
Gambar 4. 5 DFD Level 2 dari Proses Melakukan Pemodelan.....	38
Gambar 4. 6 DFD Level 2 dari Proses Evaluasi.....	39
Gambar 4. 7 Antarmuka Halaman Awal	40
Gambar 4. 8 Antarmuka Unggah Data	40
Gambar 4. 9 Antarmuka Informasi Data	41
Gambar 4. 10 Antarmuka PCA	41
Gambar 4. 11 Antarmuka Pemodelan.....	42
Gambar 4. 12 Antarmuka Evaluasi Data	42
Gambar 4. 13 Antarmuka Pedoman Penggunaan Aplikasi	43
Gambar 4. 14 Implementasi Antarmuka Sidebar	44
Gambar 4. 15 Implementasi Antarmuka Halaman Awal	45
Gambar 4. 16 Implementasi Antarmuka List Dataset	45
Gambar 4. 17 Implementasi Antarmuka Login	45
Gambar 4. 18 Implementasi Antarmuka Signup	46
Gambar 4. 19 Implementasi Antarmuka Unggah Data	46
Gambar 4. 20 Implementasi Antarmuka Logout	46
Gambar 4. 21 Implementasi Antarmuka Deskripsi Data.....	47
Gambar 4. 22 Implementasi Antarmuka Informasi Data	47
Gambar 4. 23 Implementasi Antarmuka Data Statistik.....	48
Gambar 4. 24 Implementasi Antarmuka Grafik PCA	48
Gambar 4. 25 Implementasi Antarmuka Pemodelan.....	49
Gambar 4. 26 Implementasi Antarmuka Evaluasi Data	50

Gambar 4. 27 Implementasi Antarmuka Pedoman Aplikasi	50
Gambar 4. 28 Urutan Fitur Percobaan Dataset Pima (Diabetes) pada Aplikasi Analisis Performa Isolation Forest Berbasis Konsep CAPBL	52
Gambar 4. 29 Halaman awal aplikasi	52
Gambar 4. 30 Halaman List Dataset.....	53
Gambar 4. 31 Deskripsi Data Dataset Pima	53
Gambar 4. 32 Tabel Dataset Diabetes	54
Gambar 4. 33 Data Statistik Dataset Diabetes.....	54
Gambar 4. 34 PCA 2 Dimensi Dataset Diabetes	55
Gambar 4. 35 Pengaturan Parameter Algoritma Isolation Forest.....	56
Gambar 4. 36 Hasil Evaluasi dengan $n_estimator$ 100, $contamination$ 0,341, max_sample 256, dan $test\ size$ 0,3.....	57
Gambar 4. 37 Deskripsi Dataset Cardio	59
Gambar 4. 38 Tabel Dataset Cardio	60
Gambar 4. 39 Data Statistik Dataset Cardio.....	60
Gambar 4. 40 PCA 2 Dimensi Dataset Cardio	61
Gambar 4. 41 Hasil Evaluasi dengan $n_estimator$ 100, $contamination$ 0.095, max_sample 256, dan $test\ size$ 0.3.....	61
Gambar 4. 42 Deskripsi Data Dataset Satellite	63
Gambar 4. 43 Tabel Dataset Satellite	64
Gambar 4. 44 Data Statistic Dataset Satellite.....	64
Gambar 4. 45 PCA 2 Dimensi Dataset Satellite.....	65
Gambar 4. 46 Hasil Evaluasi dengan $n_estimator$ 100, $contamination$ 0.315, max_sample 256, dan $test\ size$ 0,3.....	65
Gambar 6. 1 Hasil Perbaikan Exception Handling Unggah Data.....	77
Gambar 6. 2 Hasil Perbaikan Exception Handling Pengaturan Parameter.....	77

DAFTAR ISTILAH

No.	Istilah	Keterangan
1.	CRISP-DM	<i>Cross-Industry Standard Process for Data Mining</i> , merupakan standarisasi proses <i>data mining</i> sebagai strategi pemecahan masalah secara umum dari bisnis atau unit penelitian
2.	PBL	<i>Problem Based Learning</i> , merupakan pembelajaran yang berpusat pada pembelajar tentang suatu subjek sebagai bahan pembelajaran melalui pengalaman pemecahan masalah
3.	PBLAsF	<i>Problem Based Learning Assessment Framework</i> , merupakan kerangka penilaian yang digunakan dalam pembuatan CAPBLAT
4.	CAPBL	<i>Computer Assisted Problem Based Learning</i> , merupakan pendekatan <i>Problem Based Learning</i> dengan memanfaatkan komputer dalam pembelajaran
5.	CAPBLAT	<i>Computer Assisted Problem Based Learning Assessment Tool</i> , merupakan <i>tool</i> yang dirancang untuk membantu dosen melakukan proses pengajaran PBL
6.	DFD	<i>Data Flow Diagram</i> , merupakan metode yang digunakan dalam pemodelan suatu sistem dengan menggambar alur dari data yang ada dalam sistem
7.	DCD	<i>Data Context Diagram</i> , merupakan nama lain DFD level 0
8.	SRS	<i>Software Requirement Specification</i> , merupakan standar yang menjelaskan tentang berbagai kebutuhan yang harus dalam membangun sebuah sistem
9.	PCA	<i>Principal Component Analysis</i> , merupakan <i>tool</i> untuk menganalisis pola dalam data dengan dimensi tinggi dengan cara mereduksi data menjadi dimensi yang lebih rendah

BAB I

PENDAHULUAN

Bab ini membahas latar belakang, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan laporan dalam penyusunan skripsi mengenai pengembangan aplikasi analisis performa *Isolation Forest* berbasis *Computer Assisted Problem Based Learning* (CAPBL).

1.1. Latar Belakang

Anomali merupakan pola sebuah data yang memiliki karakteristik yang berbeda dengan data normal (Liu, Zhou, & Ting, 2008). Untuk mengenali anomali pada sebuah data maka perlu dilakukannya identifikasi pola data menggunakan teknik deteksi anomali pada *data mining*.

Deteksi anomali penting diterapkan pada berbagai bidang aplikasi untuk mengetahui kesalahan pada sebuah sistem. Dengan mengetahui anomali tersebut maka dapat meminimalisir kesalahan tersebut. Algoritma yang dapat digunakan untuk deteksi anomali antara lain *ORCA*, *Local Outlier Factor (LOF)*, *One-Class SVM (OCSVM)*, *Random Forest (RF)*, dan *Isolation Forest*. Dari beberapa contoh algoritma tersebut, salah satu algoritma deteksi anomali yang cukup diperhitungkan saat ini adalah *Isolation Forest* (Domingues, Filippone, Michiardi, & Zouaoui, 2018).

Algoritma *Isolation Forest* merupakan algoritma yang efisien dan efektif dalam menangani deteksi anomali. Algoritma ini didasarkan pada struktur pohon biner yang membangun serangkaian *isolation tree* (iTree) untuk data yang ditetapkan melalui pengambilan sampel acak (Ding & Fei, 2013). *Isolation Forest* memiliki performa yang lebih baik dari algoritma lain dalam hal waktu eksekusi, terutama dalam dataset yang besar. Pada penelitian yang dilakukan Domingues (2018), membuktikan bahwa *Isolation Forest* lebih unggul dari segi *ROC*, *precision-recall*, *robustness*, *complexity*, dan *memory usage* dibandingkan dengan algoritma deteksi anomali yang lain.

Meskipun algoritma *Isolation Forest* memiliki banyak kelebihan, namun masih sedikit *tools* yang menyediakan algoritma tersebut. Salah satu *tools* yang menyediakan algoritma *Isolation Forest* yaitu *Scikit Learn*. Namun, penggunaan *tools* tertentu seperti *Scikit Learn*

membutuhkan waktu dan pengalaman yang cukup untuk mampu menguasai fitur-fitur pada *tools* tersebut (Özmen & Altun, 2014). Hal ini menyulitkan pengguna yang tidak menguasai bahasa pemrograman dalam menggunakan algoritma *Isolation Forest*.

Dengan demikian, pada penelitian ini dilakukan pengembangan sebuah aplikasi analisis performa *Isolation Forest* sebagai pendeteksi anomali yang digunakan untuk membantu pengguna dalam mempelajari serta mengamati performa dari algoritma *Isolation Forest* dengan disertai contoh dataset, pengaturan parameter *isolation forest*, visualisasi dan hasil evaluasi.

Aplikasi analisis performa *Isolation Forest* sebagai pendeteksi anomali ini dikembangkan dengan konsep *Computer Assisted Problem Based Learning* (CAPBL) berdasarkan proses bisnis yang mengacu pada metodologi *Cross-Industry Standard Process for Data Mining* (CRISP-DM). Tujuan utama penggunaan CAPBL adalah memberikan kepraktisan bantuan kepada pengguna pada proses pembelajaran. Menurut Rajeswari (2015), model CAPBL terbukti efektif dalam meningkatkan pemikiran keterampilan siswa dalam mengembangkan tiga komponen keahlian yaitu pemikiran yang kritis, kreatif, dan keterampilan dalam memecahkan masalah. Namun, untuk menerapkan CAPBL diperlukan pemahaman terhadap proses bisnis dalam mendeteksi anomali menggunakan *Isolation Forest*. Sedangkan, *Isolation Forest* sendiri merupakan salah satu algoritma dalam kajian deteksi anomali di bidang *data mining*. Sehingga pada penelitian ini, digunakan metodologi CRISP-DM sebagai proses bisnis untuk memandu tahap-tahap pengerjaan *data mining* secara terstruktur.

1.2. Rumusan Masalah

Berdasarkan pada latar belakang yang telah dijelaskan, maka dapat dirumuskan suatu permasalahan yaitu bagaimana membuat sebuah aplikasi analisis performa *Isolation Forest* berbasis CAPBL berdasarkan proses bisnis yang mengacu pada metodologi CRISP-DM.

1.3. Tujuan dan Manfaat

Tujuan yang ingin dicapai dari pengerjaan skripsi ini adalah membangun sebuah aplikasi untuk memfasilitasi pengguna dalam menggunakan, mempelajari serta mengamati performa *Isolation Forest*. Aplikasi tersebut merupakan aplikasi berbasis *web* untuk

menganalisis performa algoritma *Isolation Forest* berbasis CAPBL yang dibangun berdasarkan metodologi CRISP-DM.

Manfaat yang diperoleh dari pengerjaan skripsi ini adalah membantu pengguna yang tidak menguasai pemrograman dalam menggunakan, mempelajari serta mengamati performa algoritma *Isolation Forest* melalui aplikasi yang disertai contoh dataset, pengaturan parameter *Isolation Forest*, visualisasi dan hasil evaluasi.

1.4. Ruang Lingkup

Ruang lingkup dalam skripsi ini adalah sebagai berikut:

1. Aplikasi yang dibuat merupakan aplikasi berbasis *web* dengan menggunakan HTML, CSS, bahasa pemrograman Python, dan framework Flask
2. Pengembangan aplikasi yang dilakukan menggunakan model proses *Waterfall* yang dilakukan sampai tahap *implementation and unit testing*.
3. Dataset yang digunakan merupakan *dataset* publik dengan tipe numerik, memiliki kelas 0 sebagai kelas normal dan kelas 1 sebagai kelas anomali, serta sudah dilakukan proses *cleaning*.
4. Dataset yang dapat diproses pada aplikasi ini disimpan dalam format file csv dan mat.
5. Pembelajaran dan pengamatan deteksi anomali pada aplikasi hanya menggunakan algoritma *Isolation Forest*

1.5. Sistematika Penulisan

Sistematika penulisan yang digunakan dalam skripsi ini terbagi dalam beberapa pokok bahasan, yaitu:

BAB I PENDAHULUAN

Bab ini membahas mengenai latar belakang masalah, rumusan masalah, tujuan dan manfaat, ruang lingkup, dan sistematika penulisan dalam penyusunan skripsi.

BAB II LANDASAN TEORI

Bab ini membahas tinjauan pustaka dan dasar teori yang digunakan untuk penyelesaian masalah, pelaksanaan dan penyusunan skripsi.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan tentang garis besar penyelesaian masalah skripsi, penjelasan singkat dari langkah-langkah yang dilakukan dan perhitungan manual dari algoritma yang digunakan.

BAB IV HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil dan pembahasan dari langkah-langkah yang dilalui berdasarkan metodologi penelitian dalam menyelesaikan permasalahan skripsi.

BAB V PENUTUP

Bab ini menjelaskan mengenai kesimpulan dari uraian yang telah dijabarkan pada bab-bab sebelumnya dan saran untuk pengembangan penelitian lebih lanjut.