# IMPLEMENTASI METODE ENSEMBLE PADA ALGORITMA KLASIFIKASI TERHADAP KASUS IMBALANCED DATASET

**SKRIPSI**


disusun oleh :

**Aditya Ahmad Zein**

**17.11.1401**



**PROGRAM SARJANA**
**PROGRAM STUDI INFORMATIKA**
**FAKULTAS ILMU KOMPUTER**
**UNIVERSITAS AMIKOM YOGYAKARTA**
**YOGYAKARTA**


**2021**

# IMPLEMENTASI METODE ENSEMBLE PADA ALGORITMA KLASIFIKASI TERHADAP KASUS IMBALANCED DATASET

**SKRIPSI**

untuk memenuhi sebagian persyaratan
mencapai gelar Sarjana
pada Program Studi Informatika


disusun oleh :

**Aditya Ahmad Zein**

**17.11.1401**



**PROGRAM SARJANA**
**PROGRAM STUDI INFORMATIKA**
**FAKULTAS ILMU KOMPUTER**
**UNIVERSITAS AMIKOM YOGYAKARTA**
**YOGYAKARTA**

**2021**

# PERSETUJUAN

## SKRIPSI

**IMPLEMENTASI METODE ENSEMBLE PADA ALGORITMA KLASIFIKASI TERHADAP KASUS IMBALANCED DATASET**

yang dipersiapkan dan disusun oleh

**Aditya Ahmad Zein**

**17.11.1401**

telah disetujui oleh Dosen Pembimbing Skripsi

pada tanggal 28 Februari 2021

**Dosen Pembimbing,**

**Yoga Pristyanto, S.Kom., M.Eng.**

**NIK. 190302412**

## PENGESAHAN

## SKRIPSI

**IMPLEMENTASI METODE ENSEMBLE PADA ALGORITMA KLASIFIKASI TERHADAP KASUS IMBALANCED DATASET**

yang dipersiapkan dan disusun oleh
**Aditya Ahmad Zein**

**17.11.1401**

telah dipertahankan di depan Dewan Penguji
pada tanggal

**Susunan Dewan Penguji**

| Nama Penguji | Tanda Tangan |
|---|---|
| **Windha Mega Pradnya D, M.Kom**<br>**NIK. 190302185** | |
| **Anna Baita, M.Kom**<br>**NIK. 190302290** | |
| **Yoga Pristyanto, S.Kom, M.Eng**<br>**NIK. 190302412** | |

Skripsi ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Sarjana Komputer
Tanggal

**DEKAN FAKULTAS ILMU KOMPUTER**

**Hanif Al Fatta,S.Kom., M.Kom**
**NIK. 190302096**

## PERNYATAAN

Saya yang bertandatangan dibawah ini menyatakan bahwa, skripsi ini merupakan karya saya sendiri (ASLI), dan isi dalam skripsi ini tidak terdapat karya yang pernah diajukan oleh orang lain untuk memperoleh gelar akademis di suatu institusi pendidikan tinggi manapun, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis dan/atau diterbitkan oleh orang lain, kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Segala sesuatu yang terkait dengan naskah dan karya yang telah dibuat adalah menjadi tanggungjawab saya pribadi.

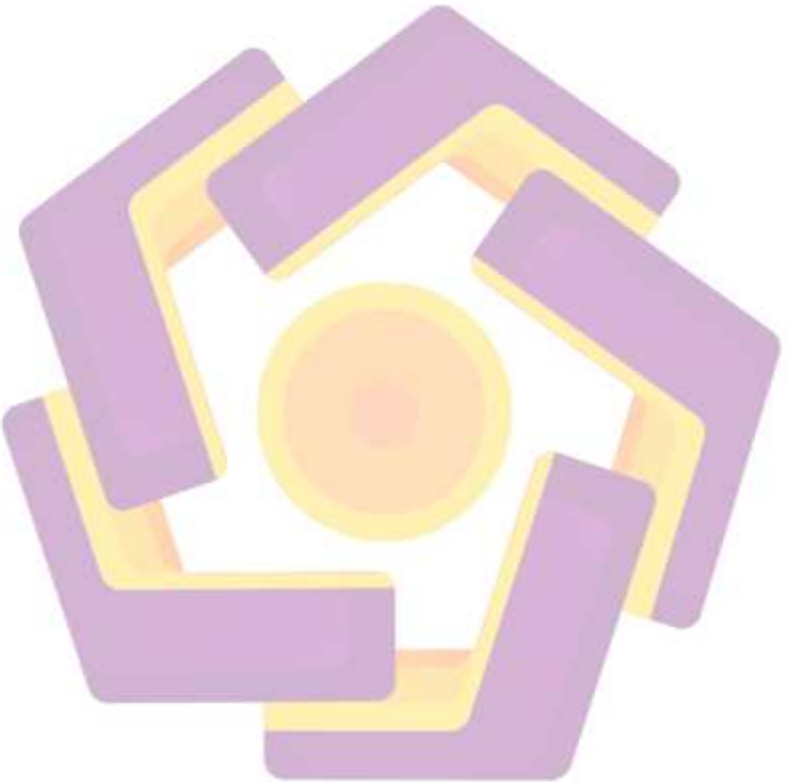
Yogyakarta, 26 Februari 2021

Aditya Ahmad Zein

NIM. 17.11.1401

## MOTTO

"Sesungguhnya Allah tidak akan mengubah nasib suatu kaum hingga mereka mengubah diri mereka sendiri," (QS. Ar-Ra'd:11).



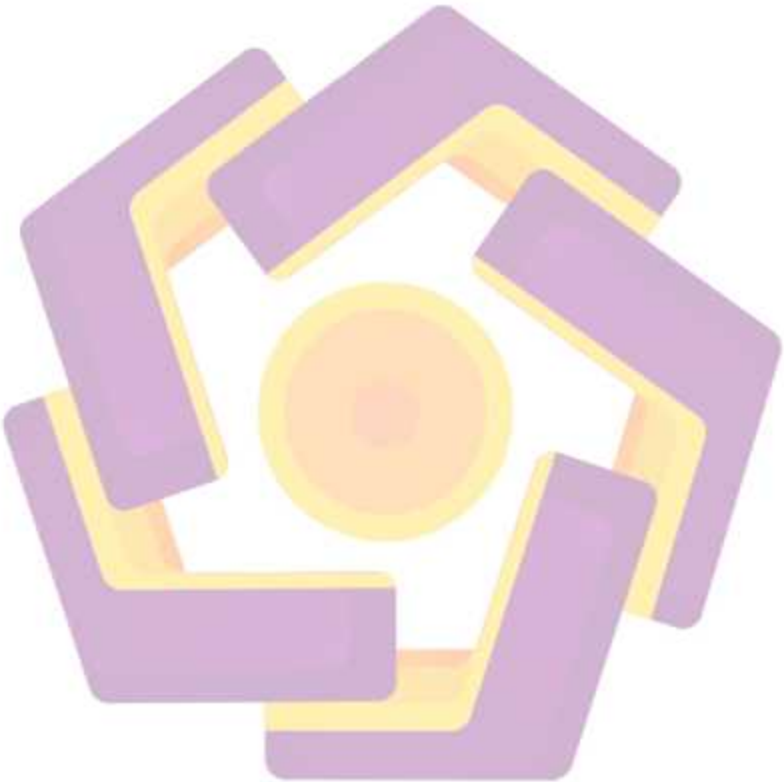
# PERSEMBAHAN

Alhamdulillah penulis panjatkan pujisyukur kepada Allah SWT atas segala rahmat, taufiq, serta hidayah-Nya, sehingga diberi kesempatan untuk dapat menyelesaikan skripsi ini dengan sebaik-baiknya dengan segala kekurangan penulis. Segala syukur penulis ucapkan kepada-Mu karena telah menghadirkan mereka yang memberikan semangat dan doa disaat menjalani proses pembuatan skripsi ini. Dengan segala kerendahan hati saya persembahkan skripsi ini kepada :

1. Kedua Orang Tua, Bapak Muh. Muslimin, M. Ag Dan Ibu Hindun Sukma Dj., S. Pd yang selalu mendoakan, memberi semangat serta motivasi supaya dapat menyelesaikan skripsi ini dengan lancar serta bermanfaat bagi semua.
2. Adikku Avillian Aufa yang tidak berhenti untuk tetap menghibur hingga memberi semangat pada saya dalam mengerjakan skripsi.
3. Bpk. Yoga Pristyanto, S.Kom, M.Eng selaku dosen pembimbing dalam skripsi ini yang tidak lelah untuk tetap membimbing dan mengingatkan penulis dari awal hingga akhir proses pembuatan skripsi.
4. Sdr. Lucky Adhikrisna Wirasakti, S.Kom. selaku teman sekaligus mentor disaat pembuatan program sedang berlangsung.
5. Dosen-dosen Universitas Amikom Yogyakarta yang telah memberikan banyak ilmu baik ilmu akademik maupun ilmu non-akademik selama kuliah.
6. Keluarga besar kelas 17-S1IF-08 yang telah bersama-sama menemani selama kuliah. Semoga silaturahmi kita tetap terjaga.
7. Keluarga besar UKM UKI Jashtis yang selalu bersama-sama berjalan di dalam dakwah selama kuliah di Amikom. Semoga kita tetap bersama sampai di surga nanti.
8. Serta orang-orang yang selalu membantu peneliti dalam mengerjakan skripsi yang tidak bisa disebut namanya satu-persatu.

Saya ucapkan terima kasih yang sebesar-besarnya untuk kalian semua. Mohon maaf jika ada salah kata atau perbuatan baik yang disengaja maupun tidak disengaja

selama ini. Sukses untuk kalian semua, semoga Allah SWT memberikan rahmat dan hidayah-Nya kepada kita semua. Dan semoga skripsi ini dapat bermanfaat dan berguna untuk kemajuan ilmu pengetahuan kedepannya.



# KATA PENGANTAR

Alhamdulillah penulis panjatkan puji syukur kepadaAllah SWT atas segala rahmat, taufiq, serta hidayah-Nya kepada penulis sehingga dapat menyelesaikan skripsi yang berjudul "Implementasi Metode Ensemble Pada Algoritma Klasifikasi Terhadap Kasus Imbalanced Dataset".

Selama prosespengerjaan skripsi ini penulis menyadari bahwa dalam proses penulisan skripsi ini banyak mengalami kendala, namun berkat bantuan, bimbingan, kerjasama dari berbagai pihak dan berjkah dari Allah SWT sehingga kendala-kendala tersebut bisa diatasi. Selanjutnya ucapa terima kasih penulis sampaikan kepada :

1. Bapak Prof. Dr. M. Suyanto, M.M selaku Rektor Universitas Amikom Yogyakarta.
2. Bpk. Yoga Pristyanto, S.Kom, M.Eng selaku dosen pembimbing yang telah memberikan banyak masukkan yang membantu membimbing dalam menyelesaikan skripsi ini.
3. Ibu Krisnawati, S.Si, M.T selaku Dekan Fakultas Ilmu Komputer Universitas Amikom Yogyakarta.
4. Bapak Sudarmawan, M.T selaku Kepala Prodi Informatika Universitas Amikom Yogyakarta.
5. Dosen Penguji (Ibu Windha Mega Pradnya D, M.Kom dan Ibu Anna Baita, M.Kom) yang telah memberikan masukkan terhadap penelitian ini.
6. Kedua orang tua dan keluarga yang selalu memberikan doa, dukungan dan semangat.
7. Keluarga Besar UKM UKI Jashtis dan Keluarga Besar 17-S1IF-08.
8. Serta semua pihak yang tidak bisa penulis sebutkan satu-persatu yang telah membantu dalam penyusunan skripsi ini.

Penulis Menyadari bahwa masih banyak terdapat kekurangan-kekurangan dalam mengerjakan skripsi ini, sehingga penulis mengharapkan adanya saran dan kritik yang membangun demi kesempurnaan skripsi ini.

Yogyakarta, 28 Februari 2021

Aditya Ahmad Zein
NIM. 17.11.1401

# DAFTAR ISI

# DAFTAR GAMBAR

# DAFTAR TABEL

# INTISARI

Algoritma klasifikasi merupakan algoritma yang sangat sering digunakan beriringan dengan kebutuhan manusia, namun peneliti sering menjumpai kendala saat menggunakan algoritma klasifikasi ini. Salah satu permasalahan yang sering sekali dijumpaioleh peneliti ialah kasus *imbalanced dataset*. Sehingga dalam penelitian ini peneliti menyarankan *ensemble method* untuk mengatasinya salah satu algoritma ensemble method yang terkenal ialah bagging. Dan disini peneliti juga akan menyarankan algoritma *balanced-bagging* untuk meningkatkan kemampuan dari bagging.

Dalam penelitian ini melibatkan tiga proses klasifikasi berbeda dengan lima dataset yang memiliki *imbalanced ratio* (IR) yang berbeda – beda yang dievaluasi berdasarkan hasil akurasi (*balanced accuracy*) dan *geometric mean*. Proses Pertama merupakan proses klasifikasi pada *Base Classifier* (tanpa Bagging) dimana mencapai performa terbaiknya pada dataset *pen_digits* dengan rata - rata hasil akurasi (92,66%) dan g-mean (91,85%). Kemudian pada proses kedua merupakan proses klasifikasi menggunakan *Base Classifier* (dengan Bagging) dengan dataset yang sama terjadi peningkatan hasil evaluasi yaitu dengan rata - rata hasil akurasi (92,68%) dan g-mean (92,31%). Namun pada beberapa dataset yang kurang berisik metode bagging mengalami penurunan kinerja sehingga dilakukan proses ketiga. Pada proses ketiga ini dilakukan klasifikasi menggunakan *Base Classifier* (dengan *Balanced-Bagging*) dimana terjadi peningkatan yang hasil evaluasi dengan dataset yang sama yaitu rata - rata hasil akurasi dan g-mean meningkat menjadi (95,90%), peningkatan hasil evaluasi ini terjadi pada setiap dataset.

Dengan peningkatan hasil evaluasi baik nilai akurasi maupun *g-mean*, menunjukkan bahwa kemampuan klasifikasi juga meningkat dan lebih peka terhadap dataset minoritas sehingga hasil peelitian ini dapat dijadikan referensi terhadap penanganan kasus *imbalanced dataset*.

**Kata Kunci**: *Bagging*, *Balanced-Bagging*, Ketidakseimbangan Kelas, Klasifikasi.

## *ABSTRACT*

*The classification algorithm is an algorithm that is very often used in conjunction with human needs, but researchers often encounter problems when using this classification algorithm. One of the problems that researchers often encounter is the case of imbalanced dataset. So that in this study the researcher suggests the ensemble method to overcome it. One of the well-known ensemble method algorithms is bagging. And here the researcher will also suggest a balanced-bagging algorithm to improve the ability of bagging.*

*In this research, it involves three different classification processes with five datasets that have different imbalanced ratio (IR) which are evaluated based on the results of balanced accuracy and geometric mean. The first process is a classification process on the Base Classifier (without Bagging) which achieves its best performance on the pen_digits dataset with an average accuracy of results (92.66%) and g-mean (91.85%). Then in the second is the classification process using the Base Classifier (with Bagging) with a dataset that has an increase in evaluation results, namely the average process result (92.68%) and the g-mean (92.31%). However, in some datasets that were less noisy, the bagging method experienced a decrease in performance so a third process was carried out. In this third process, classification is carried out using a Base Classifier (with Balanced-Bagging) where there is an increase in the evaluation results with the same dataset, namely the average value of the value and the g-mean increases to (95.90%), an increase in the results of this evaluation occurs in each data set.*

*With an increase in the results of the evaluation both the accuracy and g-mean values, it shows that the classification ability also increases and is more sensitive to minority datasets so that the results of this study can be used as a reference for handling cases of imbalanced dataset.*



***Keywords***: *Bagging, Balanced-Bagging, Class Imbalance, Decision Tree, Classification.*