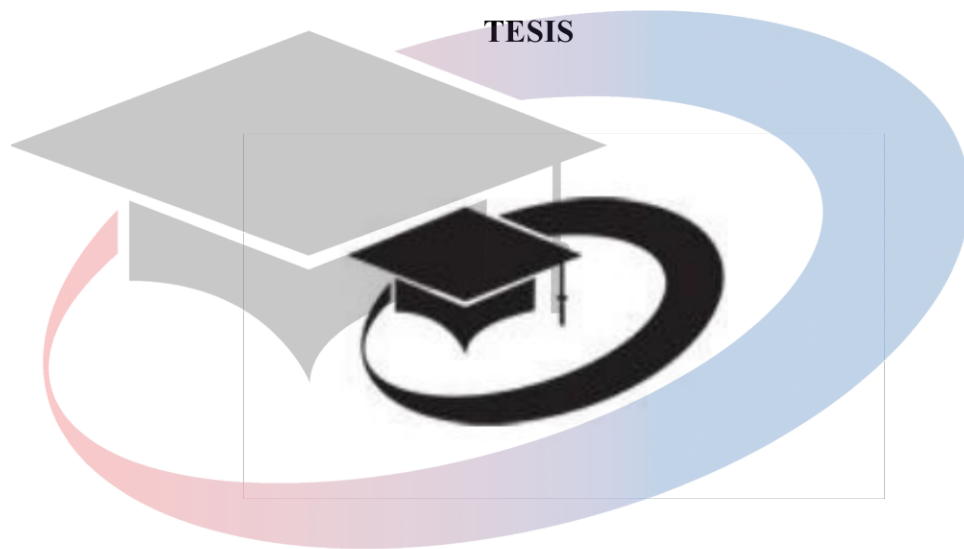


**KAJIAN KINERJA METODE SUPPORT VECTOR MACHINE
DAN NEURAL TANGENT KERNEL UNTUK
MEMPREDIKSI HASIL UJIAN SISWA**



SURYANTO
NIM. 194211045

**UNIVERSITAS
MIKROSKIL**

**PROGRAM STUDI MAGISTER TEKNOLOGI INFORMASI
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
MIKROSKIL
MEDAN
2021**

**PERFORMANCE STUDY OF SUPPORT VECTOR MACHINE
AND NEURAL TANGENT KERNEL METHODS FOR
PREDICTING STUDENT EXAMINATION
RESULTS**

TESIS



**SURYANTO
NIM. 194211045**

**UNIVERSITAS
MIKROSKIL**

**INFORMATION TECHNOLOGY MASTER PROGRAM
SEKOLAH TINGGI MANAJEMEN INFORMATIKA DAN KOMPUTER
MIKROSKIL
MEDAN
2021**

LEMBAR PERNYATAAN

Saya yang membuat pernyataan ini adalah mahasiswa Jurusan/Program Studi S-2 Magister Teknologi Informasi STMIK Mikroskil Medan dengan identitas mahasiswa sebagai berikut:

Nama : Suryanto
NIM : 194211045

Saya telah melaksanakan penelitian dan penulisan Tesis dengan judul “KAJIAN KINERJA METODE SUPPORT VECTOR MACHINE DAN NEURAL TANGENT KERNEL UNTUK MEMPREDIKSI HASIL UJIAN SISWA”, dengan ini saya menyatakan dengan sebenar-benarnya bahwa penelitian dan penulisan Tesis tersebut merupakan hasil karya saya sendiri (tidak menyuruh orang lain yang mengerjakannya) dan semua sumber, baik yang dikutip maupun dirujuk, telah saya nyatakan dengan benar. Bila di kemudian hari ternyata terbukti bahwa bukan saya yang mengerjakannya (membuatnya), maka saya bersedia dikenakan sanksi yang telah ditetapkan oleh STMIK Mikroskil Medan, termasuk pencabutan ijazah yang telah saya terima.

Selain itu, demi pengembangan ilmu pengetahuan, saya menyetujui untuk memberikan kepada STMIK Mikroskil Medan Hak Bebas Royalti Non-eksklusif (Non-exclusive Royalty Free Right) atas Tesis saya beserta perangkat yang ada (jika diperlukan). Dengan hak ini, STMIK Mikroskil Medan berhak menyimpan, mengalihmedia/formatkan, mengelola dalam bentuk pangkalan data (database), merawat, dan mempublikasikan Tesis saya, secara keseluruhan atau hanya sebagian atau hanya ringkasannya saja dalam bentuk format tercetak dan/atau elektronik, selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik hak cipta. Menyatakan juga bahwa saya akan mempertahankan hak eksklusif saya untuk menggunakan seluruh atau sebagian isi Tesis saya guna pengembangan karya di masa depan, misalnya dalam bentuk artikel, buku, ataupun perangkat lunak.

Demikian pernyataan ini saya perbuat dengan sungguh-sungguh, dalam keadaan sadar dan tanpa ada tekanan dari pihak manapun.

Medan, 04 Februari 2021

Saya yang membuat pernyataan,



Suryanto

LEMBARAN PENGESAHAN

KAJIAN KINERJA METODE SUPPORT VECTOR MACHINE DAN NEURAL
TANGENT KERNEL UNTUK MEMPREDIKSI HASIL UJIAN SISWA

TESIS

Diajukan untuk Melengkapi Persyaratan Guna

Mendapatkan Gelar Magister Strata Dua

Program Studi Magister Teknologi Informasi

Oleh:

Suryanto

NIM : 194211045

Disetujui Oleh:

Dosen Pembimbing,

Dr. Pahala Sirait, M.Kom.

Medan,

Diketahui dan Disahkan Oleh:

Ketua Program Studi

Magister Teknologi Informasi,

Dr. Ronsen Purba, M.Sc.

ABSTRAK

Data mining adalah salah satu teknik yang sangat populer dan bagus untuk prediksi. Salah satu algoritma yang sangat populer dalam data mining adalah *SVM*. Tetapi *SVM* ini memiliki kelemahan dalam menambang data-data yang besar. Untuk mengatasi hal ini terdapat metode kernel dimana salah satu metode kernel yang terbaru adalah *Neural Tangent Kernel (NTK)*.

Kombinasi penggunaan *SVM* dan *NTK* dapat digunakan dalam prediksi. Penelitian ini mengkaji metode *SVM* dikombinasikan dengan *NTK* dalam menambang *Open Learning University Dataset* yang merupakan *Dataset* pembelajaran online negara Inggris yang memiliki database besar dan tersertifikasi.

Hasil akurasi prediksi penelitian ini diukur dengan 10 Fold Validation menunjukkan hasil yang memuaskan yaitu akurasi dari Kondisi (85.23 %) lalu Performa (84.03 %) dan yang terakhir adalah Demografi (73.17 %). Saran untuk penelitian berikutnya mungkin dapat menggunakan kombinasi algoritma data mining lainnya, serta dataset pembelajaran online tersertifikasi yang lebih cocok untuk Indonesia yang diharapkan tersedia di masa depan.

Kata Kunci: *SVM*, *NTK*, Data Mining, Ujian.

UNIVERSITAS
MIKROSKIL

ABSTRACT

Data mining is one technique that is very popular and great for prediction. One of the algorithms that are very popular in data mining is SVM. But this SVM has a weakness in mining big data. To overcome this, there is a kernel method where one of the newest kernel methods is the Neural Tangent Kernel (NTK).

The combination of the use of SVM and NTK can be used in prediction. This study examines the SVM method combined with NTK in mining the Open Learning University Dataset which is the UK's online learning dataset which has a large and certified database.

The results of the prediction accuracy of this study measured by 10 Fold Validation showed satisfactory results, namely the accuracy of Condition (85.23%) then Performance (84.03%) and the last one was Demography (73.17%). Suggestions for future research might be to use a combination of other data mining algorithms, as well as certified online learning datasets that are more suitable for Indonesia which are expected to be available in the future.

Keywords: SVM, NTK, Data Mining, Examination.

UNIVERSITAS
MIKROSKIL

KATA PENGANTAR

Puji dan syukur dipanjatkan kepada Tuhan Yang Maha Esa, karena atas berkat dan rahmat-Nya, penulis dapat menyelesaikan Tesis yang berjudul “KAJIAN KINERJA METODE SUPPORT VECTOR MACHINE DAN NEURAL TANGENT KERNEL UNTUK MEMPREDIKSI HASIL UJIAN SISWA”.

Tesis ini dibuat untuk melengkapi persyaratan kurikulum pada Program Studi Magister Teknologi Informasi Strata Dua, STMIK Mikroskil Medan.

Penulis menyampaikan banyak terima kasih kepada:

1. Bapak Dr. Pahala Sirait, S.T., M.Kom. , selaku Dosen Pembimbing dan Ketua STMIK Mikroskil yang telah membimbing penulis dalam proses penyelesaian tesis ini.
2. Bapak Andri, S.Kom., M.T.I., selaku Dosen Pendamping Pembimbing yang telah membimbing penulis dalam proses penyelesaian tesis ini.
3. Bapak Dr. Ronsen Purba, M.Sc, selaku Ketua Program Studi Magister Teknologi Informasi Mikroskil Medan.
4. Bapak Ir. Erwin Setiawan Panjaitan, MMSI, Ph.D., selaku Sekretaris Program Studi Magister Teknologi Informasi Mikroskil Medan.
5. Bapak Gunawan, S.Kom., M.T.I., selaku Wakil Ketua I STMIK Mikroskil Medan.
6. Bapak atau Ibu Dosen Penguji yang telah memberikan masukan dan saran kepada penulis.
7. Bapak atau Ibu Dosen STMIK Mikroskil Medan yang telah membantu proses penulisan tesis ini.
8. Anggota keluarga, saudara-saudari, dan teman yang terus memberikan dukungan penuh kepada penulis selama proses penulisan tesis ini.
9. Semua pihak yang telah membantu penulis dalam penulisan tesis ini.

Penulis menyadari bahwa masih terdapat banyak kekurangan yang ada. Oleh sebab itu, kritik dan saran yang bersifat membangun akan sangat diterima. Akhir kata, semoga Tesis ini dapat bermanfaat bagi masyarakat. Terima kasih.

Medan, 04 Pebruari 2021

Suryanto

DAFTAR ISI

	Hal
ABSTRAK.....	iii
ABSTRACT.....	iv
DAFTAR ISI.....	vi
DAFTAR GAMBAR.....	viii
DAFTAR TABEL	x
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Masalah Penelitian	2
1.2.1 Identifikasi Masalah	2
1.2.2 Rumusan Masalah.....	3
1.3 Tujuan Penelitian dan Manfaat Penelitian.....	3
1.3.1 Tujuan Penelitian.....	3
1.3.2 Manfaat Penelitian	3
1.4 Batasan Masalah	4
1.5 Metodologi Penelitian	5
1.6 Sistematika Penulisan.....	5
BAB II KAJIAN LITERATUR	7
2.1 Tinjauan Pustaka	7
2.1.1 Data Mining.....	13
2.1.2 Pengolahan Data Mining	13
2.1.3 Konsep Prediksi	14
2.1.4 Pendekatan Dalam Data Mining.....	15
2.1.5 Algoritma Supervised Learning	16
2.1.6. Algoritma Unsupervised Learning.....	17
2.1.7 Tahapan Dalam Data Mining.....	17
2.1.8 SVM	19
2.1.9 Neural Tangen Kernel (<i>NTK</i>).....	22
2.1.10 Cross-validation (CV).....	24
2.1.11 Confusion Matrix.....	26
2.2 Penelitian Terdahulu	27

2.3	Kerangka Konsep/Pola Pikir Pemecahan Masalah	30
2.4	Hipotesis	34
BAB III METODOLOGI PENELITIAN		35
3.1.	Analisis Masalah	35
3.2	Metode Penelitian	37
3.2.1	<i>Dataset Selection</i>	38
3.2.2	Preprocessing	38
3.2.3	Mengolah <i>Dataset</i> dengan <i>SVM</i> dan <i>NTK</i>	41
3.2.4.	Pengujian Metode Cross Validation.....	43
3.2.5	Analisis hasil menggunakan confusion matrix	45
3.3.	Data yang Digunakan.....	48
3.4.	Alat Penelitian	49
3.5	Jadwal Penelitian.....	51
BAB IV HASIL DAN PEMBAHASAN		52
4.1	Hasil.....	52
4.1.1.	Data Preprocessing (Persiapan data)	52
4.1.2.	Pengkonversian data.....	56
4.1.3	<i>Dataset Demographics</i>	57
4.1.4	<i>Dataset Condition</i>	58
4.1.5	<i>Dataset Performa</i>	58
4.1.6.	Pengolahan Data	59
4.1.7.	Pengujian Data	64
4.2.	Pembahasan.....	67
4.2.1	<i>Data Preprocessing</i>	67
4.2.2.	Pengolahan Data	69
4.2.3.	Pengujian Data	70
BAB V KESIMPULAN DAN SARAN		73
5.1	Kesimpulan.....	73
5.2.	Saran	73
DAFTAR PUSTAKA		75
DAFTAR RIWAYAT HIDUP		79

DAFTAR GAMBAR

Gambar II-1 Soft margin dalam <i>SVM</i>	21
Gambar II-2 Contoh Gambar <i>NTK</i>	24
Gambar II-3 Tabel Confusion Matrix	26
Gambar II-4 Kerangka Konsep Pemecahan Masalah	31
Gambar III- 1 Diagram Metodologi Penelitian	37
Gambar III- 2 Diagram Proses Pengolahan Data	40
Gambar III- 3 Algoritma <i>SVM</i> (Sembiring, 2007)	41
Gambar III- 4 Rumus training pada <i>NTK</i>	42
Gambar III- 5 Ilustrasi K-Fold Cross Validation	45
Gambar III- 6 Confusion Matrix	45
Gambar III- 7 Algoritma Pengerjaan Akurasi	47
Gambar IV- 1 Hasil Pengkoneksian Struktur Dataset Oulad	52
Gambar IV- 2 Hasil Tampilan Potongan isi File	53
Gambar IV- 3 Hasil olahan Dataset awal	54
Gambar IV- 4 Sampel Hasil Konversi Dataset <i>SVM</i> Demographics	57
Gambar IV- 5 Sampel Hasil Konversi Sampel Dataset <i>SVM</i> Condition	58
Gambar IV- 6 Sampel Hasil Dataset <i>SVM</i> Performance	59
Gambar IV- 7 Hasil Rancangan pada Rapid Miner	60
Gambar IV- 8 Hasil Pengaturan Parameter <i>SVM</i> dan Kernel	60
Gambar IV- 9 Grafik Dot Dataset Demographic	61
Gambar IV- 10 Grafik Garis Dataset Demographic	62
Gambar IV- 11 Grafik Dot Dataset Condition	62
Gambar IV- 12 Grafik Garis Dataset Condition	63
Gambar IV- 13 Grafik Dot Dataset Performance	63
Gambar IV- 14 Grafik Garis Dataset Performance	64
Gambar IV- 15 Hasil Perancangan Proses Cross Validation	64
Gambar IV- 16 Tampilan dalam Operator Cross Validation	65
Gambar IV- 17 Sampel Data Training dan Testing	65
Gambar IV- 18 Hasil Pengujian Dataset Demographic	66

Gambar IV- 19 Hasil Pengujian Dataset Condition.....	66
Gambar IV- 20 Hasil Pengujian Dataset Performance	66
Gambar IV- 21 Hasil Penelitian Terdahulu	71



UNIVERSITAS MIKROSKIL

DAFTAR TABEL

Tabel II-1 Ringkasan Penelitian Terdahulu	28
Tabel III-1 Penyelesaian Confusion Matrix.....	46
Tabel III-2 Skema Data.....	49
Tabel III-3 Jadwal Penelitian	51
Tabel IV- 1 Tabel Pengkonversian Data.....	56
Tabel IV- 2 Hasil Cross Validation.....	67
Tabel IV- 3 Tabel Akurasi	71



UNIVERSITAS
MIKROSKIL