

SKRIPSI
IDENTIFIKASI KEPAKARAN DOSEN BIDANG
KOMPUTER DI UNIVERSITAS NEGERI JAKARTA
MENGGUNAKAN LONG SHORT-TERM MEMORY DAN
BIDIRECTIONAL LONG SHORT-TERM MEMORY



BAMBANG SETIAWAN MAULUDIN
1512622011

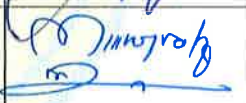
PROGRAM STUDI
PENDIDIKAN TEKNIK INFORMATIKA DAN
KOMPUTER
FAKULTAS TEKNIK
UNIVERSITAS NEGERI JAKARTA
2026

LEMBAR PENGESAHAN TUGAS AKHIR
BENTUK : SKRIPSI

Nama Mahasiswa : Bambang Setiawan Mauludin
NIM : 1512622011
Program Studi/Fakultas : Pendidikan Teknik Informatika dan Komputer / Teknik
Judul Penelitian : Identifikasi Kepakaran Dosen Bidang Komputer di Universitas
Negeri Jakarta Menggunakan Long Short-Term Memory dan
Bidirectional Long Short-Term Memory

ini telah dipertahankan dihadapan dewan penguji dan dinyatakan **Lulus** / ~~Tidak Lulus~~)*
Ujian Tugas Akhir oleh Tim Penguji pada Tanggal 9 Juni 2026

Tim Penguji,

Ketua Penguji	ZE. Ferdi Fauzan Putra, S.Pd., M.Pd.T. NIDN. 0003029004	
Anggota Penguji (Penguji Ahli 1)	Via Tuhamah Fauziastuti, S. Si, M.Ed. NIDN. 0010019110	
Anggota Penguji (Penguji Ahli 2)	Fitrah Izul Falaq, M.Pd NIDN. 0729019801	
Anggota Penguji (Pembimbing 1)	Dr. Widodo, S.Kom., M.Kom. NIDN. 0025037206	
Anggota Penguji (Pembimbing 2)	Murien Nugraheni, S.T., M.Cs NIDN. 0511108701	

Mengetahui,
Dekan Fakultas Teknik



Prof. Dr. Neneng Siti Silfi Ambarwati, S.Si., Apt., M.Si.
NIDN. 0029027204

Koordinator Program Studi
Pendidikan Teknik Informatika dan Komputer



Muchammad Ficky Duskarnaen, M.Sc
NIDN. 0024097304

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa:

1. Skripsi ini merupakan karya asli dan belum pernah diajukan untuk mendapatkan gelar akademik sarjana, baik di Universitas Negeri Jakarta maupun di perguruan tinggi lain.
2. Skripsi ini belum dipublikasikan, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan dicantumkan dalam daftar pustaka.
3. Pernyataan ini saya buat dengan sesungguhnya dan apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran, maka saya bersedia menerima sanksi akademik berupa pencabutan gelar yang telah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Universitas Negeri Jakarta.

Jakarta, 03 Juni 2026

Yang membuat pernyataan



Bambang Setiawan Mauludin

1512622011



KEMENTERIAN PENDIDIKAN TINGGI, SAINS DAN TEKNOLOGI
UNIVERSITAS NEGERI JAKARTA
PERPUSTAKAAN DAN KEARSIPAN

Jalan Rawamangun Muka Jakarta 13220

Telepon/Faksimili: 021-4894221

Laman: lib.unj.ac.id

**LEMBAR PERNYATAAN PERSETUJUAN PUBLIKASI
KARYA ILMIAH UNTUK KEPENTINGAN AKADEMIS**

Sebagai sivitas akademika Universitas Negeri Jakarta, yang bertanda tangan di bawah ini, saya:

Nama : Bambang Setiawan Mauludin
NIM : 1512622011
Fakultas/Prodi : Fakultas Teknik/ Pendidikan Teknik Informatika dan Komputer
Alamat email : b50901019@gmail.com

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Perpustakaan dan Kearsipan Universitas Negeri Jakarta, Hak Bebas Royalti Non-Eksklusif atas karya ilmiah:

☒ Skripsi ☐ Tesis ☐ Disertasi ☐ Lain-lain (.....)

yang berjudul :

Identifikasi Kepakaran Dosen Bidang Komputer di Universitas Negeri Jakarta

Menggunakan Long Short-Term Memory dan Bidirectional Long Short-Term Memory

Dengan Hak Bebas Royalti Non-Eksklusif ini Perpustakaan dan Kearsipan Universitas Negeri Jakarta berhak menyimpan, mengalihmediakan, 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 Universitas Negeri Jakarta, segala bentuk tuntutan hukum yang timbul atas pelanggaran Hak Cipta dalam karya ilmiah saya ini.

Demikian pernyataan ini saya buat dengan sebenarnya.

Jakarta, 3 Agustus 2026

Penulis

(Bambang Setiawan Mauludin)

KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala rahmat dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Identifikasi Kepakaran Dosen Bidang Komputer di Universitas Negeri Jakarta Menggunakan *Long Short-Term Memory* dan *Bidirectional Long Short-Term Memory*”.

Skripsi ini disusun sebagai salah satu syarat untuk memperoleh gelar Sarjana Pendidikan pada Program Studi Pendidikan Teknik Informatika dan Komputer, Fakultas Teknik, Universitas Negeri Jakarta. Penulis menyadari sepenuhnya bahwa perjalanan penyusunan skripsi ini tidak terlepas dari doa, bimbingan, dukungan, serta semangat dari berbagai pihak. Oleh karena itu, dengan rasa syukur dan hormat, penulis ingin menyampaikan terima kasih yang sebesar-besarnya kepada:

1. Bapak dan Mimi tercinta, atas limpahan kasih sayang yang tak terhingga, kesabaran dalam mendidik, serta doa yang senantiasa menjadi penguat dan penyemangat utama penulis dalam melewati setiap proses perkuliahan hingga tahap ini.
2. Muhammad Ficky Duskarnaen, M.Sc., selaku Koordinator Program Studi Pendidikan Teknik Informatika dan Komputer, atas segala dukungan dan arahannya selama masa studi.
3. Dr. Widodo, S.Kom., M.Kom., selaku Dosen Pembimbing I, atas kebaikan hati beliau meluangkan waktu untuk berbagi ilmu, memberikan arahan, serta masukan yang sangat berharga selama proses penelitian dan penyusunan skripsi ini.
4. Murien Nugraheni, S.T., M.Cs., selaku Dosen Pembimbing II, atas kesabaran, bantuan, serta dorongan semangat yang senantiasa diberikan dalam mendampingi penulis menyelesaikan skripsi ini.
5. Saudara-saudara terkasih penulis: Susi Suciyati, Nur Azizah, Fadiah Agustin, dan Zahratul Salamah. Terima kasih karena senantiasa hadir, memberikan kasih sayang, dan menjadi kawan diskusi yang selalu menguatkan penulis dalam berbagai keadaan.

6. Bang Mada, selaku kakak tingkat yang senantiasa terbuka untuk diajak bertukar pikiran, layaknya dosen pembimbing ketiga bagi penulis.
7. Rekan-rekan di Pustikom: Bang Bebe, Bang Alfin, Bang Zikri, Bang Rama, Bang Danan, Adani, Axel, Dhika, dan Wafi, atas pengalaman, ruang belajar, dan banyak hal baru yang telah diajarkan kepada penulis.
8. Teman-teman seperjuangan: Stevano, Ibnu, Raul, Fauzi, dan Miftah, atas doa, semangat, dan masukan berharga yang selalu menyertai perjalanan perkuliahan penulis.
9. Sahabat-sahabat baik: Iqbal, Semi, Bintang, Miftah, Nizhar, Raka, Yuan, Davin, dan Farhan, atas kebersamaan, canda tawa, serta dukungan moral yang luar biasa.
10. Seluruh pihak yang turut membantu dalam penyelesaian skripsi ini, baik secara langsung maupun tidak langsung, yang tidak dapat penulis sebutkan satu per satu tanpa mengurangi rasa hormat.

Penulis menyadari bahwa skripsi ini masih jauh dari kata sempurna dan memiliki berbagai keterbatasan. Oleh karena itu, kritik dan saran yang membangun sangat penulis harapkan demi perbaikan di masa mendatang. Akhir kata, semoga skripsi ini dapat memberikan manfaat dan kontribusi yang berarti bagi pengembangan ilmu pengetahuan.

Jakarta, 06 Juni 2026



Bambang Setiawan Mauludin

NIM.1512622011

IDENTIFIKASI KEPAKARAN DOSEN BIDANG KOMPUTER DI UNIVERSITAS NEGERI JAKARTA MENGGUNAKAN LONG SHORT-TERM MEMORY DAN BIDIRECTIONAL LONG SHORT- TERM MEMORY

Bambang Setiawan Mauludin

Dosen Pembimbing: Dr. Widodo, S.Kom., M.Kom. dan Murien
Nugraheni, S.T., M.Cs.

ABSTRAK

Pemetaan kepakaran dosen di bidang komputer Universitas Negeri Jakarta (UNJ) saat ini belum terotomatisasi secara terstruktur. Penelitian ini bertujuan menganalisis dan membandingkan performa model *deep learning Long Short-Term Memory* (LSTM) dan *Bidirectional LSTM* (BiLSTM) untuk mengklasifikasikan publikasi ilmiah dosen berdasarkan standar *ACM Computing Classification System* (ACM CCS) 2020. Data yang digunakan berjumlah 149 dokumen publikasi terindeks Scopus dari dosen program studi Ilmu Komputer, PTIK, dan STI. Metode yang diterapkan meliputi pra-pemrosesan teks untuk mereduksi *noise* dan pemodelan topik *Latent Dirichlet Allocation* (LDA) guna mendukung proses pelabelan data secara hibrida. Fitur teks diekstraksi dan dibandingkan menggunakan representasi *Word Embedding* jenis *Word2Vec* dan *FastText*. Hasil evaluasi pada 129 dokumen valid (setelah penyaringan kategori NON-CC2020) menunjukkan bahwa *FastText* lebih superior dalam menangani variasi istilah teknis komputasi (*Out-of-Vocabulary*) dibandingkan *Word2Vec*. Kombinasi *FastText* dengan model LSTM searah menghasilkan performa paling optimal dengan tingkat akurasi 84,62% dan *macro F1-score* 67,73%. Sebaliknya, arsitektur BiLSTM justru menunjukkan penurunan akurasi (80,77%) akibat kejenuhan representasi (*over-parameterization*) pada data latih yang terbatas. Model terbaik ini secara efektif berhasil memetakan tren kepakaran dosen UNJ, yang saat ini didominasi oleh *Data Science* (22,15%), *Information Technology* (18,79%), dan *Computer Science* (17,45%). Hasil ini diharapkan dapat menjadi landasan penting untuk pengembangan institusi di masa depan.

Kata Kunci: Identifikasi Kepakaran, Klasifikasi Teks, *ACM CCS*, LSTM, *FastText*.

**Identification of Computer Field Lecturers' Expertise at Universitas
Negeri Jakarta Using Long Short-Term Memory and Bidirectional Long
Short-Term Memory**

Bambang Setiawan Mauludin

**Supervisors: Dr. Widodo, S.Kom., M.Kom. and Murien Nugraheni, S.T.,
M.Cs.**

ABSTRACT

The mapping of lecturers' expertise in the computer field at Universitas Negeri Jakarta (UNJ) was not structurally automated. This study aimed to analyze and compare the performance of Long Short-Term Memory (LSTM) and Bidirectional LSTM (BiLSTM) deep learning models in classifying lecturers' scientific publications based on the 2020 ACM Computing Classification System (ACM CCS) standard. The data used consisted of 149 Scopus-indexed publication documents from lecturers in the Computer Science, PTIK, and STI study programs. The applied methods included text preprocessing to reduce noise and Latent Dirichlet Allocation (LDA) topic modeling to support the hybrid data labeling process. Text features were extracted and compared using Word Embedding representations, specifically Word2Vec and FastText. The evaluation results on 129 valid documents (after filtering out the NON-CC2020 category) showed that FastText was superior in handling variations of computational technical terms (Out-of-Vocabulary) compared to Word2Vec. The combination of FastText with the unidirectional LSTM model yielded the most optimal performance with an accuracy rate of 84.62% and a macro F1-score of 67.73%. In contrast, the BiLSTM architecture exhibited a decrease in accuracy (80.77%) due to representational saturation (over-parameterization) on the limited training data. This best model effectively mapped the trends in UNJ lecturers' expertise, which was dominated by Data Science (22.15%), Information Technology (18.79%), and Computer Science (17.45%). These results were expected to serve as an important foundation for future institutional development.

Keywords: Expertise Identification, Text Classification, ACM CCS, LSTM, FastText.

DAFTAR ISI

LEMBAR PENGESAHAN UJIAN SKRIPSI	II
LEMBAR PERNYATAAN	III
KATA PENGANTAR	IV
ABSTRAK	VI
ABSTRACT	VII
DAFTAR ISI	VIII
DAFTAR GAMBAR	XI
DAFTAR TABEL	XII
DAFTAR LAMPIRAN	XIV
BAB I PENDAHULUAN	1
1.1 Latar Belakang Masalah	1
1.2 Identifikasi Masalah	4
1.3 Pembatasan Masalah	5
1.4 Rumusan Masalah	6
1.5 Tujuan Penelitian	6
1.6 Manfaat Penelitian	6
BAB II TINJAUAN PUSTAKA	7
2.1 Kerangka Teoritik	7
2.1.1 Publikasi Ilmiah	7
2.1.2 Program Studi Bidang Komputer	8
2.1.3 Text Preprocessing	9
2.1.4 Natural Language Processing (NLP)	11
2.1.5 <i>Topic Modelling</i>	12
2.1.6 <i>Latent Dirichlet Allocation</i>	13
2.1.7 <i>Coherence Score</i>	15
2.1.8 ACM Computing Classification System (ACM CCS)	17
2.1.9 Pendekatan Hibrida <i>Rule-Based Expert System</i>	18
2.1.10 Word Embedding	20
2.1.11 Deep learning	23
2.1.12 Long Short-Term Memory	24
2.1.13 Bidirectional Long Short-Term Memory	26

2.1.14 Confussion Matirix	27
2.2 Penelitian Relevan	29
2.3 Kerangka Berpikir.....	37
BAB III METODE PENELITIAN	40
3.1 Tempat dan Waktu Penelitian.....	40
3.1.1 Alat Penelitian.....	40
3.2 Alat dan Bahan Penelitian.....	41
3.2.2 Bahan Penelitian.....	41
3.3 Diagram Alir Penelitian	43
3.4 Prosedur Penelitian	44
3.4.1 Pengumpulan Publikasi Ilmiah	44
3.4.2 Text Preprocessing	45
3.4.3 Pemodelan Topik Menggunakan LDA.....	47
3.4.4 Pemetaan Topik dan Pelabelan Data	48
3.4.5 Representasi Teks (<i>Word Embedding</i>).....	54
3.4.6 Desain Skenario Eksperimen dan Skema Pembagian Data	55
3.4.7 Identifikasi Kepakaran dengan LSTM dan BiLSTM.....	56
3.5 Teknik Pengumpulan data.....	59
3.6 Metode Evaluasi Data	60
BAB IV HASIL DAN PEMBAHASAN.....	62
4.1 Hasil Penelitian	62
4.1.1 Pengumpulan Data Publikasi Ilmiah.....	62
4.1.2 Hasil <i>Text Preprocessing</i>	64
4.1.3 Pemodelan Topik (LDA).....	68
4.1.4 Pelabelan Bidang Kepakaran Kurikulum ACM CC2020	73
4.1.5 Analisis <i>Crosstab</i> Topik LDA dan Label CC2020	77
4.1.6 Ekstraksi Fitur (<i>Word Embedding</i>).....	78
4.1.7 Hasil Pelatihan Model LSTM dan BiLSTM	81
4.2 Analisis hasil Penelitian.....	84
4.2.1 Evaluasi Skenario 1 (Model LSTM + Word2Vec).....	84
4.2.2 Evaluasi Skenario 2 (Model LSTM + FastText).....	85
4.2.3 Evaluasi Skenario 3 (Model BiLSTM + Word2Vec)	87

4.2.4	Evaluasi Skenario 4 (Model BiLSTM + FastText)	88
4.2.5	Perbandingan Kinerja Keseluruhan Skenario	89
4.3	Pembahasan.....	91
4.3.1	Analisis Pengaruh Arsitektur dan Representasi Kata.....	91
4.3.2	Identifikasi Peta Kepakaran Dosen Bidang Komputer Universitas Negeri Jakarta.....	93
BAB V KESIMPULAN DAN SARAN		98
5.1	Kesimpulan	98
5.2	Saran	99
DAFTAR PUSTAKA.....		101
LAMPIRAN.....		108
DAFTAR RIWAYAT HIDUP		117



DAFTAR GAMBAR

Gambar 2. 1 Plate Notation Latent Dirichlet Allocation (Blei et al., 2009)..	14
Gambar 2. 2 Arsitektur CBOW dan Skip-gram pada Word2Vec (Mikolov et al., 2013).....	21
Gambar 2. 3 Ilustrasi Subword pada FastText (Bojanowski et al., 2017).....	22
Gambar 2. 4 Arsitektur Sel Long Short-Term Memory (LSTM).....	25
Gambar 2. 5 Arsitektur Bidirectional LSTM (BiLSTM)	27
Gambar 2. 6 Kerangka Berpikir	39
Gambar 3. 1 Diagram Alir Penelitian.....	43
Gambar 4. 1 Grafik Perbandingan Jumlah Kata Sebelum dan Sesudah Prapemrosesan	67
Gambar 4. 2 Grafik Evaluasi Coherence Score (C_v) dan Log Perplexity Pemodelan Topik LDA.....	69
Gambar 4. 3 Visualisasi Top 15 Kata Kunci Dominan untuk 12 Topik Terpilih	72
Gambar 4. 4 Diagram Lingkaran Distribusi Bidang Kepakaran Dosen Berdasarkan Taksonomi ACM CC2020.....	74
Gambar 4. 5 Heatmap Distribusi Jumlah Dokumen Hasil Tabulasi Silang Topik LDA dan Label CC2020	78
Gambar 4. 6 Grafik Kurva Pelatihan Skenario 1 (LSTM + Word2Vec).....	82
Gambar 4. 7 Grafik Kurva Pelatihan Skenario 2 (LSTM + FastText).....	83
Gambar 4. 8 Grafik Kurva Pelatihan Skenario 3 (BiLSTM + Word2Vec) ...	83
Gambar 4. 9 Grafik Kurva Pelatihan Skenario 4 (BiLSTM + FastText)	83
Gambar 4. 10 Confusion Matrix Skenario 1	84
Gambar 4. 11 Confusion Matrix Skenario 2	86
Gambar 4. 12 Confusion Matrix Skenario 3	87
Gambar 4. 13 Confusion Matrix Skenario 4	88
Gambar 4. 14 Grafik Batang Komparatif Performa Metrik Klasifikasi Antarmodel.....	90

DAFTAR TABEL

Tabel 2. 1 Confusion Matrix	28
Tabel 2. 2 Penelitian Relevan.....	32
Tabel 3. 1 Perangkat keras	40
Tabel 3. 2 Perangkat Lunak.....	41
Tabel 3. 3 Contoh Hasil Preprocessing	46
Tabel 3. 4 Distribusi Label Kelas ACM Computing Curricula 2020	53
Tabel 3. 5 Skenario Pengujian Hyperparameter Model LSTM/BiLSTM.....	58
Tabel 4. 1 Distribusi Pengumpulan Data Publikasi Ilmiah per Program Studi	63
Tabel 4. 2 Langkah-Langkah dan Hasil Transformasi Pra-pemrosesan Teks	66
Tabel 4. 3 Rangkuman Hasil Eksperimen Grid Search Pemodelan Topik LDA	68
Tabel 4. 4 Rincian 12 Topik Terpilih Hasil Pemodelan LDA	70
Tabel 4. 5 Distribusi Hasil Pelabelan Bidang Kepakaran Dosen Universitas Negeri Jakarta	73
Tabel 4. 6 Distribusi Jumlah Dokumen Hasil Tabulasi Silang Topik LDA dan Label CC2020	77
Tabel 4. 7 Spesifikasi parameter pelatihan Word2Vec dan FastText	79
Tabel 4. 8 Perbandingan Nilai Cosine Similarity pada Kata Kunci Acuan Komputasi	80
Tabel 4. 9 Konfigurasi Hyperparameter untuk Arsitektur LSTM dan BiLSTM	81
Tabel 4. 10 Metrik Evaluasi (Precision, Recall, F1-Score) per Kelas ACM CCS untuk Skenario 1	84
Tabel 4. 11 Metrik Evaluasi per Kelas untuk Skenario 2.....	86
Tabel 4. 12 Metrik Evaluasi per Kelas untuk Skenario 3.....	87
Tabel 4. 13 Metrik Evaluasi per Kelas untuk Skenario 4.....	89
Tabel 4. 14 Rekapitulasi Perbandingan Performa (Accuracy dan Macro F1-Score) Keempat Model	90

Tabel 4. 15 Pemetaan Tiga Besar Bidang Kepakaran Dosen Rumpun Komputasi UNJ.....	93
Tabel 4. 16 Sampel Hasil Prediksi Bidang Kepakaran ACM CC2020 Berdasarkan Judul Artikel Dosen.....	95



DAFTAR LAMPIRAN

Lampiran 1 Publikasi Ilmiah Dosen Terindeks scopus	108
Lampiran 2 Kode Program.....	109
Lampiran 3 Surat Tugas Dosen Pembimbing	110
Lampiran 4 Surat Pernyataan Penyelesaian Mata Kuliah Wajib.....	111
Lampiran 5 Lembar Persetujuan Dosen Pembimbing	112
Lampiran 6 Lembar Konsultasi Skripsi (Dosen Pembimbing I).....	113
Lampiran 7 Lembar Konsultasi Skripsi (Dosen Pembimbing I).....	115

