# *Re-Fake:* Klasifikasi Akun Palsu di Sosial Media *Online* menggunakan Algoritma RNN

## *(Re-Fake: Classification of Fake Accounts on Online Social Media using the RNN Algorithm)*

**Putra Wanda[1*], Marselina Endah Hiswati[2], Mohammad Diqi[3], Romana Herlinda[4]**

[1,2,3] Prodi S-1 Informatika, Universitas Respati Yogyakarta, Indonesia
*E-mail: putra.wanda@respati.ac.id, marsel.endah@respati.ac.id, diqi@respati.ac.id*
[3] Informatika, Universitas Respati Yogyakarta
*E-mail: 17220022@respati.ac.id*

Abstrak— ***Online Social Network (OSN) adalah aplikasi Social Media yang memungkinkan komunikasi publik dan berbagi informasi. Namun, akun palsu di OSN dapat menyebarkan informasi palsu dengan sumber yang tidak diketahui. Sebuah cara yang cukup sulit untuk mendeteksi akun berbahaya dalam sistem OSN dalam ekosistem yang besar. Keberadaan akun palsu atau akun tidak dikenal di OSN dapat menjadi masalah serius dalam menjaga privasi data. Berbagai komunitas telah mengusulkan banyak teknik untuk menangani akun palsu di OSN, misalnya teknik Black-White listing hingga pendekatan pembelajaran. Oleh karena itu, dalam penelitian ini kami mengusulkan model klasifikasi menggunakan RNN untuk mendeteksi akun palsu secara akurat dan efektif. Kami melakukan penelitian ini dalam beberapa langkah, meliputi tahap mengumpulkan dataset, pra-pemrosesan, ekstraksi, melatih model RNN. Berdasarkan hasil eksperimen, model yang kami usulkan dapat menghasilkan akurasi yang lebih tinggi daripada model pembelajaran konvensional.***

Kata Kunci— **Klasifikasi, Akun Palsu, *Recurrent Neural Network, Deep Learning*.**

Abstract— ***Online Social Network (OSN) is an application that enables public communication and information sharing. However, fake accounts on OSN can spread false information with unknown sources. It is a challenging task to detect malicious accounts in a large OSN system. The existence of fake accounts or unknown accounts on OSN can be a serious problem in maintaining data privacy. Various communities have proposed many techniques to deal with fake accounts on OSN, including rules-based black-and-white techniques to learning approaches. Therefore, in this study we propose a classification model using RNN to detect fake accounts accurately and effectively. We carried out this research in several steps, including collecting the dataset, pre-processing, extraction, training our model using RNN. Based on the experimental results, our proposed model can produce higher accuracy than conventional learning models.***

Keywords— **Classification, Fake Account, Recurrent Neural Network, Deep Learning.**

** Penulis Korespondensi (Putra Wanda)*
*Email: putra.wanda@respati.ac.id*

p-ISSN 2086-5805
e-ISSN 2808-2540

# I. PENDAHULUAN

Munculnya OSN telah memicu menyebarnya berbagai informasi palsu di kalangan masyarakat pengguna OSN[1]. Akun palsu di OSN dapat menyebarkan berita palsu. Akun anonim, fiktif, dan akun samar lainnya digunakan oleh individu untuk mengekspresikan diri, memanfaatkan media sosial, dan melakukan aktivitas lain di dunia maya tanpa mengungkapkan identitas aslinya kepada orang lain. Akibat dari informasi palsu yang disebarluaskan dapat menyesatkan masyarakat yang mengandalkan OSN sebagai media untuk memperoleh informasi, dan dapat misinformasi misinformasi di kalangan pengguna OSN [2].

Berkembangnya aplikasi Sosial Media juga dapat membuka peluang yang luas bagi orang-orang tertentu untuk melakukan penipuan dengan menggunakan akun palsu [3]. Salah satu kecurangan dalam OSN adalah penyalahgunaan identitas seseorang atau instansi yang nantinya akan digunakan untuk melakukan kejahatan seperti jual beli barang atau berbisnis. Penipuan ini dapat menimbulkan kesalahpahaman antara pengguna dalam jual beli barang dan berbisnis sehingga banyak orang atau pengguna harus berhati-hati dalam melakukan transaksi online di OSN [4].

Dalam lingkungan OSN, komentar spam dianggap tidak relevan secara kontekstual, sering ditemukan di OSN [5]. Sebagai contoh komentar spam pada halaman OSN Instagram *public figure*, komentar spam tersebut dapat berupa postingan yang tidak ada kaitannya dengan postingan dan status pada halaman OSN yang bersangkutan, seperti mengirimkan informasi yang tidak diinginkan oleh pengguna [6]. Aktivitas spam memang mengganggu karena dapat menimbulkan informasi yang menyesatkan dan mengganggu alur diskusi dalam status hingga kesulitan mencari informasi. Apalagi komentator spam di OSN dilakukan oleh akun palsu yang identitasnya tidak diketahui [7].

*Cyberbullying* juga merupakan salah satu yang patut mendapat perhatian dari dampak akun palsu pada OSN sebagai tindakan pelecehan menggunakan teknologi [8]. Aktivitas tersebut biasanya dilakukan di OSN dalam bentuk komentar jahat, posting gambar, atau video yang dimaksudkan untuk menyakiti atau mempermalukan orang lain. Hal ini berdampak besar baik bagi masyarakat maupun pengguna OSN karena dapat menimbulkan kepanikan yang berujung pada kematian. Selain itu, *cyberbullying* juga dapat menyebabkan penurunan mental bagi pengguna yang menjadi korbannya [9].

Untuk menangani masalah akun berbahaya, beberapa makalah mengusulkan pendekatan pembelajaran (*learning*) untuk menghasilkan akurasi yang lebih baik. Makalah yang dibahas adalah penerapan *deep learning* menggunakan *GitSec* pada komunitas *developer online* seperti *GitHub*, hal ini dilakukan karena komunitas ini terbuka untuk umum sehingga dapat membuatnya rentan terhadap berbagai jenis serangan jahat yang dilakukan oleh akun palsu seperti spam, penipuan, dan penyebaran informasi palsu [10]. Selain itu, algoritma CNN dan LSTM juga diterapkan untuk mengklasifikasikan akun palsu di OSN. *Deep learning* ke OSN dilakukan karena informasi atau berita yang disebarkan oleh akun anonim beredar sangat cepat, sehingga membuat pengguna kesulitan mencari atau mengetahui kebenaran berita atau informasi tersebut [11].

Oleh karena itu, untuk mengatasi masalah di atas, kami mengusulkan *deep learning* untuk membangun model klasifikasi akun palsu pada media sosial. Penelitian ini memiliki kontribusi pada studi akun palsu sebagai berikut:

1. Membangun model klasifikasi baru untuk mengidentifikasi akun palsu di OSN. Untuk mengukur kinerja model, kami melakukan proses pelatihan dan pengujian serta menghitung akurasi dan kerugian.

2. Melakukan metrik evaluasi dalam mendeteksi akun palsu di OSN dan menyajikan grafik untuk membuktikan kualitas model. Dalam penelitian ini, kami menggunakan *dataset* akun palsu untuk menghasilkan model klasifikasi. Model mengolah dataset sebagai input dan menggunakan model tersebut untuk memudahkan masyarakat dalam mendeteksi akun palsu.

3. Dengan menggunakan algoritma pembelajaran, kami menghadirkan model klasifikasi yang efisien menggunakan metode *deep learning*. Alih-alih menggunakan metode konvensional, model yang diusulkan dapat mendeteksi akun palsu di OSN secara akurat.

Organisasi: Sisa dari makalah ini akan ditulis sebagai berikut: Bagian II menggali lebih jauh penelitian terkait. Bagian III menguraikan bagaimana mendefinisikan masalah studi ini. Bagian IV membahas pengaturan eksperimental, termasuk pendekatan pembelajaran fitur, kumpulan data, dan pra-pemrosesan data, sementara Bagian V memberikan temuan studi dan analisis ekstensif. Terakhir, bagian VI merangkum temuan dan mengidentifikasi beberapa masalah yang belum terpecahkan dalam penelitian kategorisasi akun palsu.

## II. LANDASAN TEORI DAN TINJAUAN PUSTAKA

Saat ini, beberapa artikel mengusulkan berbagai penelitian, seperti penelitian menggunakan algoritma *Naïve Bayes* memperoleh hasil dengan akurasi yang wajar hanya dengan pra-pemrosesan *dataset* menggunakan teknik diskrit pada fitur yang dipilih [12]. Studi lain membahas algoritma SVM menggunakan fitur yang lebih sedikit tetapi dapat menghasilkan secara akurat [13]. Sebuah makalah juga mengeksplorasi algoritma *bagging* untuk menghasilkan hasil yang akurat, dengan kesalahan yang rendah [14].

Pengklasifikasian akun palsu menggunakan algoritma DL mendapatkan hasil yang akurat dan signifikan untuk menciptakan proteksi di OSN [15]. Untuk mendeteksi akun palsu, peneliti lain menggunakan SVM dan *neural network*. Metode SVM mencapai skor F1 86%, dan jaringan saraf mencapai skor F1 95% [16]. Studi lain menggunakan metode OWL dan SWRL untuk mengklasifikasikan akun palsu. Sebaliknya, hasil penelitian ini mampu secara tepat mengidentifikasi akun yang salah dengan akurasi (97%) dan hasil *spam* atau *bot follower* palsu dengan tingkat akurasi (94,9%) [17].

Sebuah makalah juga mengeksplorasi klasifikasi menggunakan metode siku. Menerapkan prinsip analisis komponen untuk mengklasifikasikan akun palsu mendapatkan hasil yang menunjukkan akurasi tingkat keberhasilan 99,6% dan tingkat kegagalan 0% [18]. Studi lain juga membahas algoritma *Sybil Walk* yang dapat menghasilkan hasil yang lebih akurat daripada metode random walk berbasis *Sybil Walk* yang ada mencapai rasio positif palsu 1,3% dan tingkat negatif 17,3% [19]. Dalam penelitian lain, klasifikasi dapat mendeteksi dan secara akurat mengidentifikasi akun palsu pada OSN yang disalahgunakan untuk kejahatan [20].

Pendekatan konvensional mengusulkan metode SVM dan CNB yang dapat menghasilkan hasil dengan *dataset Facebook*. SVM menunjukkan akurasi 97%, dan CNB menunjukkan akurasi 95% untuk mengidentifikasi akun palsu berdasarkan BOW [21]. Makalah lain juga mengusulkan metode *Mandatory Unique Identification* Model (CUIM) untuk membantu meningkatkan keamanan dan secara efektif menangani pemegang akun palsu [22]. Pada penelitian lain, klasifikasi akun palsu menggunakan metode SVM, RF, dan NN. SVM menunjukkan kinerja akurasi prediksi yang lebih tinggi dalam mengklasifikasikan akun palsu daripada RF dan NN [23].

p-ISSN 2086-5805
e-ISSN 2808-2540

Beberapa makalah mengeksplorasi klasifikasi akun palsu menggunakan CNN. Penelitian ini dapat menghasilkan hasil yang akurat dengan ROC 0.9500 - 0.9590 dan AUC 0.9547 [24]. Artikel lain menyajikan teknik CNN berdasarkan tujuan pin yang cocok dan nama papan. Diperoleh 886444 pin dalam 3920 akun, 1503 akun palsu berisi 345000 pin dengan akurasi klasifikasi 90,25% [25]. Makalah yang membahas tentang klasifikasi akun palsu menghasilkan akurasi yang baik dalam mendeteksi dan mengidentifikasi akun palsu [26].

Pada artikel lain, model klasifikasi menggunakan CNN dan RNN dapat mengklasifikasikan akun palsu di Twitter dengan akurasi 82% [27]. Sebuah penelitian menggunakan model deep learning hybrid yang menghubungkan CNN dan RNN untuk membangun model klasifikasi akun palsu, dan dapat mencapai hasil yang signifikan [28]. Menggunakan metode pembelajaran untuk membuat skema perlindungan adalah salah satu tren penelitian paling intensif dalam keamanan jaringan. Ini membuka peluang luas untuk mengatasi kendala metode pembelajaran mesin tradisional. Dalam algoritma pembelajaran mesin tradisional, fitur diekstraksi oleh manusia [29]. Oleh karena itu, kami mengusulkan model baru untuk menyelesaikan deteksi akun palsu di OSN dengan melatih fitur besar menggunakan algoritma RNN. Dengan demikian, tidak hanya untuk mendeteksi akun palsu tetapi juga untuk membantu pengguna dalam mendeteksi akun palsu secara akurat. Dengan demikian, model ini dapat digunakan sebagai bahan untuk penelitian lebih lanjut tentang klasifikasi akun palsu.

## III. MODEL YANG DIUSULKAN

Bagian ini akan memberikan pernyataan formal tentang masalah percobaan dan beberapa ide yang dibahas dalam artikel ini.

### *A. Definisi Masalah*

Penelitian kami berfokus pada pendeteksian akun palsu di OSN dengan mungumpulkan dataset akun palsu di OSN. Pada penelitian ini, kami membuat model klasifikasi akun palsu menggunakan algoritma RNN dengan membagi data menjadi dua kelas, dan informasi mewakili fitur (a) dan bias (b). Proses klasifikasi tidak menggunakan data pada fungsi yang memiliki parameter. Kemudian fungsi akan menghitung bobot setiap fitur dalam vektor dengan mengalikannya dengan parameter. Dengan demikian, persamaan satu dapat ditulis ulang sebagai persamaan 2, di mana a adalah elemen-i dari vektor a, yang memiliki jangkauan.

TABEL I
VARIABEL DAN INFORMASINYA

| Variabel | Informasi |
|---|---|
| $x$ | *Input* |
| $s$ | *Hidden Layer* |
| $y$ | *Output* |
| $x_t$ | *Input* pada waktu |
| $x_{tt-1}$ | *Input* pada waktu sebelumnya |
| $s_t$ | *Hidden State* pada waktu |
| $s_{tt-1}$ | *Hidden State* pada waktu sebelumnya |
| $y_t$ | *Output* pada *t-time* |
| $y_{t-1}$ | *Output* pada waktu sebelumnya |
| $f$ | Hubungan iterasi dengan fungsi aktivasi |
| $w$ | Parameter matriks dan vektor |
| $a$ | Fitur |
| $b$ | Bias |

$$f(a) = a.w + b \quad (1)$$

$$f(a) = a_1 w_1 + a_1 w_2 + \cdots + a_N w_N + b \quad (2)$$

Dalam penelitian kami, kami menggunakan regresi fungsional untuk klasifikasi akun palsu di OSN. Karena fungsi regresi ini akan menghasilkan nilai yang konstan, maka digunakan ambang batas, atau akan diberikan batas nilai tertentu. Misalnya, $f(a) > threshold$ jika dimasukkan ke dalam kelas pertama dan sebaliknya $f(a) \leq threshold$ dimasukkan ke dalam kelas kedua Pendekatan *threshold* dilakukan dengan mengubah proses menjadi -1, dan 1 sebagai *output* (persamaan 4), dimana -1 mewakili *input* yang diklasifikasikan ke dalam kelas pertama dan nilai 1 menunjukkan *input* yang diklasifikasikan ke dalam kelas kedua, menggunakan fungsi tanda (Persamaan 3).

$$sgn(a) = \begin{cases} -1 \; if \; a < 0 \\ 0 \; if \; a = 0 \\ 1 \; if \; > 0 \end{cases} \quad (3)$$

$$Output = sgn(f(a)) \quad (4)$$

## *B. Metode yang Diusulkan*

Penelitian ini menggunakan algoritma RNN untuk membangun model klasifikasi *deep learning* untuk mendeteksi akun jahat di OSN. Untuk membangun model kami, kami mengimplementasikan algoritma RNN untuk membuat model klasifikasi deep learning. Pemodelan RNN merupakan cara yang efektif untuk mengatasi akun palsu pada OSN karena kemampuan untuk memprosesnya disebut berulang-ulang dengan hasil yang dapat menangani variabel *input* dan *output* dengan panjang yang bervariasi [30].

Dalam penelitian ini, algoritma RNN digunakan untuk menyimpan informasi dari masa lalu dengan melakukan *looping* pada arsitekturnya, yang secara otomatis menyimpan informasi pada state yang tersimpan sebelumnya. Pengulangan atau *looping* pada RNN adalah mengambil nilai input $x$ kemudian memasukkannya ke dalam RNN yang berisi nilai *hidden layer*, yang akan di*update* setiap kali RNN membaca *input* dan *output* baru. Dalam komputasi menggunakan RNN, ada *f* fungsi *f* akan tergantung pada bobot *w*. Berat w akan menerima nilai negara baru dari lapisan tersembunyi minus 1, yang menjadi masukan dalam $x_t$ keadaandan disimpan dalam $s_t$ (tersembunyi negara). Proses ini dilakukan dengan menggunakan persamaan 5.

$$s_t = f_w(s_{t-1}, x_t) \quad (5)$$

Nilai $x$ dimasukkan ke dalam fungsi aktivasi *f* dan yang sama *w* berat di setiap perhitungan. Sederhananya, sebuah $w_{xs}$ matriks bobot dikalikan dengan $x_t$ input dan $w_{ss}$ lainnya matriks bobot dikalikan dengan nilai lapisan tersembunyi sebelumnya atau $s_{t-1}$.kedua matriks ditambahkan. Jika terdapat data *nonlinier*, maka penjumlahan kedua matriks dikalikan dengan salju, seperti pada persamaan 6.

$$s_t = tanh(W_{ss} s_{t-1} + W_{xs} x_t) \quad (6)$$

Arsitektur RNN menghasilkan beberapa $y_t$setiap saat karena ada matriks bobot lain *w* dari lapisan tersembunyi $w_s$sehingga mengubah beberapa nilai $y\ yang$ terlihat pada persamaan 7.

$$y_t = W_{sy} s_t \quad (7)$$

p-ISSN 2086-5805
e-ISSN 2808-2540

### *C. Ide Utama*

Dalam beberapa tahun terakhir, *deep learning* telah mencapai kinerja yang baik dalam menyelesaikan permasalahan *Computer Vision*. Model pembelajaran dalam ini diterapkan secara luas untuk membangun skema perlindungan untuk jaringan komputer, termasuk adopsi algoritma RNN dalam penelitian keamanan jaringan. Ide utama dari penelitian ini adalah membangun model klasifikasi dengan menggunakan algoritma RNN untuk mengklasifikasikan akun palsu di OSN berdasarkan pengikut, nama, dan tanggal akun terdaftar. Penggunaan algoritma RNN untuk membangun model klasifikasi dapat mencapai hasil yang signifikan dan akurat karena kemampuan untuk memprosesnya disebut berulang [31].

### *D. Dataset*

Dalam penelitian ini, kami menggunakan *dataset* akun palsu yang diperoleh dari situs *github.com*. Ini memiliki beberapa fitur, pengikut, nama, dan tanggal yang direkam. Kami membagi dataset menjadi dua bagian dalam proses pelatihan dan pengujian, yaitu data pelatihan dan pengujian data [32]. Untuk melakukan penelitian ini, kami mengumpulkan *dataset* akun palsu yang diadopsi dalam penelitian ini berjumlah 2.818 sampel yang digunakan untuk penelitian membangun model klasifikasi *deep learning* agar lebih mudah untuk mendeteksi akun palsu. Untuk membangun model kami, kami memisahkan 80% sebagai data pelatihan dan 20% sebagai data pengujian. Pada tabel 1 berisi jumlah *dataset training* dan *testing*.

TABEL II
MENJELASKAN DATASET YANG DIGUNAKAN UNTUK MELATIH DAN MENGUJI MODEL

| Dataset Label | fitur OSN | |
|---|---|---|
| | Pelatihan (80%) | Pengujian 20%) |
| Akun Palsu | 314 | 78 |
| Akun | 314 | 79 |

### *E. Pra-Pemrosesan Data*

Pada tahap ini, kami melakukan pra-pemrosesan dengan menggunakan vektorisasi proses untuk mengubah bentuk data yang sebelumnya tidak terstruktur menjadi terstruktur. Kemudian pada tahap *pre-processing*, *dataset* yang digunakan adalah *dataset* akun palsu sebanyak 785 data, yang akan divektorkan menggunakan metode *tokenizer*, dengan normalisasi fitur dan seleksi fitur. Terakhir, metode *tokenizer* digunakan dalam proses *pre-processing* untuk memudahkan algoritma RNN dalam menerima data input [33].

### *F. Metode Klasifikasi*

Pada tahap pertama, kami mengumpulkan dataset dengan label normal dan palsu sebelum proses pelatihan. Kemudian masuk ke tahap *pre-processing*, yaitu tahap dimana akan dilakukan vektorisasi *dataset*. Tahap vektorisasi ini dilakukan untuk memudahkan dalam membangun model klasifikasi (pelatihan) menggunakan algoritma RNN. Kami memisahkan *dataset* menjadi dua bagian, 80% pelatihan, dan 20% pengujian, untuk melatih model kami. Pertama, kita harus melatih model kita untuk membangun model klasifikasi akun palsu yang optimal menggunakan *dataset* pelatihan. Kemudian model tersebut akan diuji menggunakan *dataset testing* untuk mengetahui performansi model tersebut. Terakhir, untuk mengoptimalkan nilai pelatihan, kami juga menyetel beberapa *hyperparameter* untuk mendapatkan model klasifikasi yang optimal.

# IV. IMPLEMENTASI MODEL DAN PEMBAHASAN

## *A. Uji Klasifikasi*

Untuk mengklasifikasikan akun palsu, kami menguji model klasifikasi dengan parameter pembelajaran yang berbeda. Secara umum, kami memeriksa beberapa pengoptimal pelatihan, salah satunya adalah Adam. Kami juga menerapkan konsep variabel lokal untuk menghitung fungsi penurunan berat badan dengan berbagai kecepatan pembelajaran.

Dalam percobaan ini, model kami menghasilkan akurasi dengan menyesuaikan berbagai *hyperparameter* untuk kinerja tertinggi. Dalam proses pelatihan dan pengujian, kami menyesuaikan *epoch* = 50, ukuran *batch* = 64, dan kecepatan pembelajaran 0,2. Pada uji coba yang telah dilakukan, model dapat mengklasifikasikan dengan tingkat akurasi 81,0%. Tabel III menggambarkan kinerja RNN, terutama dalam pelatihan dan pengujian. Tabel IV Hasil klasifikasi dengan berbagai fungsi *optimizer.*

TABEL III
KINERJA RNN DALAM PELATIHAN PENGUJIAN

| *Hyperparameter* | *Optimizer* | Pelatihan Akurasi | Pengujian Akurasi |
|---|---|---|---|
| *Epoch* = 50 | Adam | 0,8678 | 0 9363 |
| Ir = 0,0002 ukuran bets 64 | RMSProp | 0,8742 | 09108 |
| kecepatan belajar 0,2 | SGD | 0,8806 | 0 8854 |

TABEL IV
HASIL KLASIFIKASI DENGAN BERBAGAI FUNGSI OPTIMIZER

| Hyperparameter | Optimizer | Pelatihan Rugi | Pengujian Loss |
|---|---|---|---|
| Epoch = 50 | Adam | 0.6807 | 0.6781 |
| Ir = 0.0002 ukuran batch 64 | RMSProp | 0.3792 | 0.3038 |
| kecepatan belajar 0.2 | SGD | 0.3659 | 0.3563 |

## *A. Metrik Evaluasi*

Penelitian ini juga menghitung *Recall, Precision, F1 Score*, dan *Confusion Matrix* untuk mengevaluasi model kinerja (CF) dalam proses metrik evaluasi. Penelitian ini menggunakan *Recall* untuk mengidentifikasi pecahan dari kumpulan data yang termasuk dalam kategori berbahaya dan melabelinya. *Precision* mendefinisikan tingkat kumpulan data, yang diklasifikasikan sebagai berbahaya tetapi tidak termasuk data berbahaya. Pada tahap akhir, kami menghitung keseimbangan antara *Precision* dan *Recall* untuk menemukan semua akun dengan tautan berbahaya sambil menoleransi Presisi yang buruk. Untuk menyelesaikan kalkulasi balance, skor F1 diperlukan untuk mewakili kombinasi akurasi dan *Recall* terbaik. Nilai *Recall, Precision*, dan *F1 Score* ditunjukkan pada Tabel V.

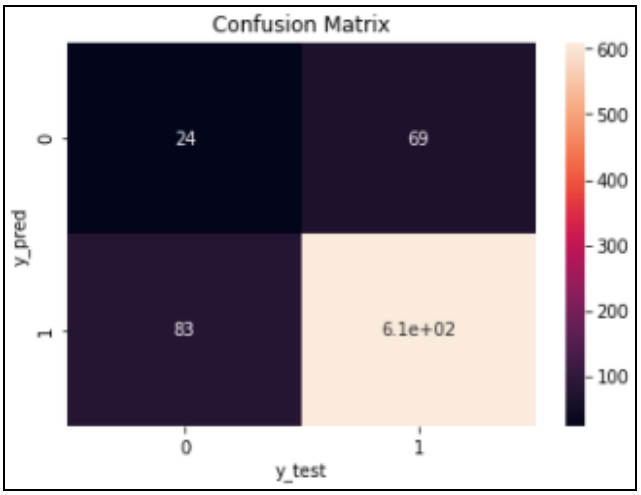
p-ISSN 2086-5805
e-ISSN 2808-2540

TABEL V
NILAI RECALL, PRESCISION, DAN F1 SCORE

| Model Klasifikasi | Precision | Recall | F1-score |
|---|---|---|---|
| False | 0.26 | 0.22 | 0.24 |
| True | 0.88 | 0.90 | 0.89 |
| RNN | | | 0.81 |
| Macro avg | 0.57 | 0.56 | 0.56 |
| Wighted avg | 0.88 | 0.81 | 0.80 |

### *B. Confusion Matrix*

Model yang diusulkan dapat memperoleh nilai akurasi terbaik untuk mewakili tingkat deteksi maksimum dalam mengkategorikan akun normal dan palsu, berdasarkan perhitungan CM pada Gambar 1. Pendekatan kami, khususnya, memberikan tidak hanya akurasi yang tinggi tetapi juga meningkatkan kinerja grafis.




Gambar 1. Algoritma Confusion Matrix RNN

Makalah ini juga menyajikan tabel *Confusion Matrix* yang menggambarkan kinerja model pada data uji yang diketahui. *Confusion matrix*, khususnya, berisi informasi tentang *True Positive* (TP), *False Positive* (FP), *True Negative* (TN), dan *False Negative* (FN). Ini sangat membantu karena hasil kategorisasi seringkali tidak cukup terwakili dalam satu angka. Gambar 1 menunjukkan evaluasi metrik dengan matriks konfusi (CM) menggunakan algoritma CNN. Kita dapat melihat angka yang diprediksi oleh pengklasifikasi kita dari matriks konfusi, secara terpisah untuk kedua kelas. Dalam matriks konfusi, model yang diusulkan diperoleh TP = 24, TN = 692, FP = 69, dan FN = 83.

## V. KESIMPULAN

Deteksi akun palsu di OSN merupakan langkah penting dalam mengkategorikan keberadaan akun palsu pada OSN. Pendekatan saat ini banyak menggunakan pembelajaran mesin tradisional untuk mengatasi masalah tersebut. Namun, pendekatan menggunakan Teknik pembelajran konvensional ini mahal dan memakan waktu. Oleh karena itu, penelitian ini mengajukan model kategorisasi akun palsu menggunakan algoritma RNN untuk meningkatkan kinerja grafik. Saat

menerapkan model yang diusulkan, kami menemukan bahwa teknik RNN menghasilkan akurasi yang lebih baik dan *loss* lebih sedikit. Selain itu, algoritma RNN untuk mebangun model klasifikasi memerlukan kemampuan pemrosesan perangkat keras yang bagus untuk mengakselerasi proses *training model*.

Berdasarkan hasil experiment, klasifikasi akun palsu menggunakan RNN dapat mencapai akurasi yang tinggi dengan kerugian yang kecil. Nilai akurasi RNN yang kami dapatkan adalah 0,81 sedangkan nilai *F1-score* adalah 0,80, *Recall* adalah 0,81, dan *Precision* adalah 0,88. Model yang kami usulkan tidak hanya dapat menghasilkan akurasi yang lebih tinggi tetapi juga meningkatkan kinerja *neural network* yang dubangun. Akibatnya, kami menyarankan bahwa model klasifikasi berbasis RNN mungkin merupakan pendekatan potensial untuk mengidentifikasi akun OSN palsu.

Penelitian lebih lanjut tentang klasifikasi akun palsu dapat mengembangkan model yang lebih bagus dengan mengubah arsitektur jaringan *neural network* yang digunakan, yang dapat diintegrasikan dengan pendekatan baru seperti algoritma CNN di masa depan dikombinasikan dengan teknik baru seperti membuat *regulator* baru untuk melatih jaringan.

## UCAPAN TERIMA KASIH

Ucapan terima kasih atas terbitnya naskah ini pada Seminar Nasional Sains Teknologi dan Inovasi Indonesia 2021. Penelitian ini dilaksanakan di bawah Departemen Informatika Universitas Respati Yogyakarta, Indonesia.

## REFERENSI

[1] Wilson Ceron, Mathias Felipe and Marcos G. Quiles, -Fake news agenda in the era of COVID-19: Identifying trends through fact-checking content, Online Social Networks and Media, Vol. 21, pp 100-116, 2021.
[2] Sy.Yuliani, Shahrin Sahib, et al., -Hoax News Classification using Machine Learning Algorithms, International Journal of Engineering and Advanced Technology, Vol.9, No.2, pp 2249-8958, 2019.
[3] Tahereh Pourhabibi et al., -Fraud detection: A systematic literature review of graph-based anomaly detection approaches, Decision Support Systems, Vol.133, No.113303, 2020.
[4] Arun Vishwanath, -Habitual Facebook Use and its Impact on Getting Deceived on Social Media, Journal of Computer-Mediated Communication Vol. 20, No. 1, pp 83-98, 2015
[5] Shreyas Aiyara, Nisha P Shetty, -N-Gram Assisted Youtube Spam Comment Detection, International Conference on Computational Intelligence and Data Science (ICCIDS), Vol.132, pp 174-182, 2018.
[6] Zulfikar Aloma, Barbara Carminatib and Elena Ferrari, -A deep learning model for Twitter spam detection, Vol.18, No. 100079, 2020.
[7] Zakia Zaman, Sadia Sharmin, -Spam Detection in Social Media Employing Machine Learning Tool for Text Mining, IEEE International Conference on Signal-Image Technology and Internet Based Sys, Vol.978, No.1, pp 4283-5386, 2018.
[8] W.Akram, R.Kumar, -A Study on Positive and Negative Effects of Social Media on Society, International Journal of Computer Sciences and Engineering, Vol.5, No.10, pp 2347-2693, 2017.
[9] Batol et all, *-Cyberbullying Detection:A Survey On Mltilingual Techniques*, IEEE European Modelling Symposium, pp 2473-3539, 2017.
[10] Qingyuan Gong, et al.- Detecting Malicious Accounts in Online Developer Communities Using Deep Learning, In The 28th ACM International Conference on Information and Knowledge Management (CIKM '19), 2019.
[11] Muhammad Umer, et al.- Fake News Stance Detection Using Deep Learning Architecture (CNN-LSTM) IEEE, Vol.8, 2020.
[12] Ersahin Buket et all, -Twitter Fake Account Detection, IEEE International Conference on Computer Science and Engineering, pp 388–392, 2017.
[13] Sarah Khaled, Hoda M. O. Mokhtar and Neamat El-Tazi, - Detecting Fake Accounts on Social Media, IEEE International Conference on Big Data, 2018.
[14] Saeid Sheikhi, -An Efficient Method for Detection of Fake Accounts on the Instagram Platform, Revue D Intelligence Artificielle, Vol.34, No.4, pp 429-436, 2020.
[15] Putra Wanda, Marselina Endah Hiswati , and Huang J. Jie, *-DeepOSN: Bringing deep learning as malicious detection scheme in online social network*, Journal of Information Security and Applications, 52, 2020.
[16] Fatih Cagatay Akyon , Esat Kalfaoglu, -Instagram Fake and Automated Account Detection Insagram Sahte ve Otomatik Hesap Kullanımı Tespiti, IEEE, Vol.978, No.1, pp 7281-2868, 2019.

p-ISSN 2086-5805
e-ISSN 2808-2540

[17] Mohammed Jabardi, -Twitter Fake Account Detection and Classification using Ontological Engineering and Semantic Web Rule Language, Karbala International Journal of Modern Science, Vol. 6, No. 4, pp 404-413, 2020.
[18] Mohammadreza Mohammadrezaei, Mohammad Ebrahim Shiri and Amir Masoud Rahmani, -Detection of fake accounts in social networks based on One Class Classification, The ISC Int'l Journal of Information Security, Vol.11, No.2, pp 1–12, 2019.
[19] Mohammadreza Mohammadrezaei, Mohammad Ebrahim Shiri and Amir Masoud Rahmani, -Detection of fake accounts in social networks based on One Class Classification, The ISC Int'l Journal of Information Security, Vol.11, No.2, pp 1–12, 2019.
[20] Disha Agarwal, Atrakesh Pandey, -Determining Fake Accounts on Facebook, International Journal of Management, Technology And Engineering, Vol. IX, No.IV, pp 2249-7455, 2019.
[21] Putra Wanda, Huang Jin Jie, -DeepProfile: Finding fake profile in online social network using dynamic CNN, Journal of Information Security and Applications, Vol. 52, 2020.
[22] Enas Elgeldawi et all, -Detection And Characterization Of Fake Accounts On The Pinterest Social Network, International Journal of Computer Networking, Wireless and Mobile Communications, Vol. 4, No. 3, 2250-1568, 2278-9448, 2014.
[23] Jia Jinyuan, Wang Binghui and Gong Neil Zhenqiang, -Random Walk Based Fake Account Detection in Online Social Networks, IEEE International Conference on Dependable Systems and Networks, pp 273–284, 2017.
[24] Rohit Raturi, -Machine Learning Implementation for Identifying Fake Accounts in Social Network, International Journal of Pure and Applied Mathematics, Vol. 118, No. 20, pp 4785-4797, 2018.
[25] Saumya Batham et all, -CUIM: An Approach to deal with Fake Accounts on facebook, International Conference on Emerging Research in Computing Information Communication and Applications (ERCICA-13), Elsevier, 2014.
[26] dr.K.Sreenivasa Rao, dr.G.Sreeram and dr. B.Deevena Raju, -Detecting Fake Account On Social Media Using Machine Learning Algorithms, International Journal Of Control And Automation, Vol. 13, No. 1, pp 95-100, 2020.
[27] Ajao, et al, -Fake News Identification on Twitter with Hybrid CNN and RNN Models, Proceedings of the 9th International Conference on Social Media and Society - SMSociety '18, pp 226–230, 2018.
[28] Osama et al, -Fake news detection: A hybrid CNN-RNN based deep learning approach, International Journal of Information Management Data Insights, Vol.1, No.1, 2021.
[29] Imamverdiyev, Yadigar N. and Fargana J. Abdullayeva, -Deep Learning in Cybersecurity: Challenges and Approaches, IJCWT vol.10, no.2, pp.82-105. 2020
[30] Alex Sherstinsky, -Fundamentals of Recurrent Neural Network (RNN) and Long Short-Term Memory (LSTM) network, Physica D: Nonlinear Phenomena, Vol.404, No.132306, pp 0167-2789, 2020.
[31] Buber, Ebubekir and Diri, Banu, -Web Page Classification Using RNN, Procedia Computer Science, Vol.158, pp 62-72, 2019.
[32] Hongyu Liu, et all, -CNN and RNN based payload classification methods for attack detection, Accepted Manuscript, 2018.
[33] Huang, Ji, et all, -Detecting Domain Generation Algorithms With Convolutional Neural Language Models, IEEE International Conference On Trust, pp 1360-1367, 2018.