

1. Pendahuluan

1.1 Latar Belakang

Kehadiran kosmetik di era ini memberi kemudahan khususnya bagi kaum wanita dalam membuat dirinya tampak menarik. Dalam industri kosmetik tanah air saat ini menarik pesat yang dipicu dengan gaya hidup generasi milenial yang merasakan tuntutan untuk selalu tampil menarik [1]. Pada era modern ini konsumen dapat menulis opini dan pengalaman mereka terhadap produk kecantikan secara *online* semakin meningkat. Dengan membaca *review* tersebut secara keseluruhan dapat memakan waktu tetapi pada saat hanya sedikit *review* yang dibaca mendapatkan hasil yang kurang maksimal saat mengevaluasi. *Review* mengenai produk membantu konsumen dalam mengetahui kualitas produk [2].

Dengan perkembangannya *machine learning* menjadi *deep learning* dalam klasifikasi sentimen diikuti juga dengan perkembangannya *word embedding*. *Word embedding* adalah metode *unsupervised learning* menggunakan *neural network* [3]. Analisis sentimen berkaitan dengan pengelompokan polaritas dari teks dalam suatu dokumen, kalimat, atau pernyataan. Polaritas mengacu pada teks dalam dokumen, kalimat, atau pernyataan yang memiliki nuansa positif atau negatif [4]. Analisis sentimen digunakan untuk mengambil informasi berharga dari data yang tidak terstruktur [5].

Word2vec merupakan salah satu dari algoritma *word embedding* yang mengubah setiap teks dalam kalimat menjadi vektor. Yang mempresentasikan kata ke dalam vektor yang membawa makna dari kata tersebut [6]. Efek yang dimiliki klasifikasi sentimen dalam teks bergantung pada penggalan fitur yang ringkas dan informatif dari teks yang tidak terstruktur. Melalui pembelajaran representasi, model representasi utama terhadap klasifikasi

sentimen dapat dibagi menjadi dua kategori berdasarkan informasi dan pengetahuan yang digunakan [7]. Metode ini digunakan untuk melakukan klasifikasi pada *website* Female Daily [8]. Sebagai algoritma dalam *deep learning*, *Convolutional neural network* (CNN) bekerja dengan melatih data pada dataset yang besar, menggunakan parameter yang menerima input berbentuk dimensi [9]. Dalam analisis sentiment, terdapat sebuah Teknik *deep learning* yang dikenal sebagai CNN [10].

Pada tahun 2021, Afidah et al., menyatakan penelitian ini salah satu hal yang mempengaruhi hasil dari parameter *Word2Vec* terhadap *deep learning* yaitu dimensi. Dimensi yang terlalu kecil dan dimensi yang terlalu besar dapat mempengaruhi nilai akurasi. Dimensi 100 adalah dimensi yang paling optimal dalam mempengaruhi akurasi model. Hal ini terlihat dari hasil rata-rata penelitian akurasi dimensi 100, 200, dan 300 secara berurutan sebesar 97,10%, 96,77% dan 96,73%. Oleh karena itu, dimensi 100 selaras dengan jumlah dataset 10.000 ulasan. Sedangkan dimensi yang lebih besar akan menurunkan hasil dari performa model [3].

Berikutnya pada tahun 2022, penulis Zahroni menyatakan dalam penelitian ini, telah melakukan beberapa model untuk mengetahui hasil akurasi terbaik. Dalam penelitian yang telah dilakukan telah melakukan 4 percobaan yang menggunakan dataset split 70% train 30% test, 80% train 20% test, 90% train 10% test dan dataset train dan dataset test terpisah. Menghasilkan akurasi dari model yang telah dibentuk tidak mengalami perbedaan yang signifikan dengan menggunakan train 70% test 30% menghasilkan 98,48% dan dataset terpisah dengan akurasi tertinggi 99,76% [11].

Selain itu, pada penelitian di tahun 2022, Hidayat dan Handayani menyatakan pada penelitiannya bahwa, setelah melakukan 30 skenario yang dijalankan, pada skenario ke-12 menghasilkan model terbaik untuk analisis sentimen. Model yang dilakukan menggunakan *learning rate* 0,008 dengan konfigurasi *layer Conv1D* (128,3), MaxPooling 1D, Dropout (0.5), Dense (64), Dense (32), Dropout (0.2), Dense (16), GlobalMaxPooling 1D, dan menjalani 60 epoch, dan menghasilkan hasil akurasi, yaitu 80,22% [1].

Dalam penelitian ini, melakukan penggabungan *Word2Vec* dan CNN dengan melakukan pendekatan yang kuat dalam pemrosesan teks. *Word2Vec* menyediakan representasi dari kata yang bermakna, sementara CNN mampu mengekstraksi dari representasi tersebut, yang menghasilkan nilai akurasi dalam menangani kompleksitas dan Bahasa dengan baik.

1.2 Tujuan

Tujuan dari penelitian ini adalah untuk melihat bagaimana metode Convolutional Neural Network (CNN) bekerja saat menganalisis produk kecantikan dan mengetahui bagaimana penggunaan fitur Word2Vec

memengaruhi hasil analisis. Peneliti ingin menentukan seberapa efektif integrasi Word2Vec sebagai representasi kata dalam CNN dalam mengklasifikasikan dan memahami ulasan produk kecantikan. Diharapkan bahwa penelitian ini akan memberikan pemahaman yang lebih mendalam tentang kombinasi metode pembelajaran mendalam dan teknik pengolahan bahasa alami (NLP) dalam industri kecantikan dan bagaimana hal-hal ini berdampak pada pembuatan model yang lebih cerdas dan adaptif.