Analisis Sentimen pada Media Sosial Terhadap Perkuliahan Hybrid Menggunakan Algoritma TF IDF dan K Nearest Neighbor

Ida Bagus Ketut Surya Arnawa
Institut Teknologi dan Bisnis STIKOM Bali
e-mail: arnawa@stikom-bali.ac.id

Diajukan: ...; Direvisi: ...; Diterima: ...

Abstrak

Tahun 2019 merupakan tahun yang penuh cobaan bagi masyarakat dunia dimana pada tahun tersebut menyebar Pandemi Covid-19 termasuk di Indonesia. Hampir semua sekolah dan kampus menghentikan operasionalnya. Demi meminimalisir dampak dari berhentinya kegiatan belajar mengajar maka perkuliahan jarak jauh dipilih sebagai solusi. Perkuliahan dilakukan melalui berbagai media seperti Ms.Teams, Zoom dan lain sebagainya. Berbagai kendala pun bermunculan dalam pelaksanaan perkuliahan ini mengingat ini adalah pertama kali dilakukan. Infrastruktur dan kesiapan SDM menjadi kendala utama yang dihadapi. Tidak jarang orang tua dan mahasiswa memberikan keluh kesah terkait dengan kebijakan baru ini didunia pendidikan. Saat Pandemi Covid-19 mulai melandai menjadi sebuah momentum bagi Perguruan Tinggi untuk berinovasi mencari alternatif dalam melakukan perkuliahan. Adapun alternatif yang diciptakan yaitu perkuliahan secara hybrid. Mengingat perkuliahan Hybrid ini juga pertama kali dilakukan tentunya memiliki kekurangan. Guna memperbaiki kekurangan-kekurangan yang ada maka perlu dilakukan evaluasi perkuliahan hybrid ini. Evaluasi yang digunakan yaitu berupa analisis sentimen opini publik terkait perkuliahan hybrid. Dari penelitian yang sebelumnya sudah pernah dilakukan, penelitian yang penulis lakukan memiliki perbedaan yaitu dari segi sumber data dan metode yang digunakan TF-IDF dan K-NN. Hasil dari penelitian ini yaitu sentimen positif sebesar 30%, netral 20% dan negatif 50% terhadap Perkuliahan Hybrid.

Kata kunci: *Hybrid, Text Mining, Covid-19.*

Abstract

2019 was a year full of trials for the world community where the Covid-19 Pandemic spread, including in Indonesia. Almost all schools and colleges stopped their operations. In order to minimize the impact of the cessation of teaching and learning activities, distance lectures were chosen as a solution. Lectures are conducted through various media such as Ms.Teams, Zoom and so on. Various obstacles also emerged in the implementation of this lecture, considering that this was the first time it was held. Infrastructure and readiness of human resources are the main obstacles faced. Not infrequently parents and students give complaints related to this new policy in the world of education. When the Covid-19 Pandemic began to decline, it became a momentum for Higher Education to innovate to find alternatives in conducting lectures. The alternative created is a hybrid lecture. Considering that this Hybrid lecture is also the first time it has been conducted, of course it has drawbacks. In order to correct the existing deficiencies, it is necessary to evaluate this hybrid lecture. The evaluation used is in the form of sentiment analysis of public opinion regarding hybrid lectures. From previous research that has been done, the research that the author did has differences, namely in terms of data sources and methods used by TF-IDF and K-NN. The results of this study are 30% positive sentiment, 20% neutral and 50% negative towards Hybrid Lectures.

Keywords*: Hybrid, Text Mining, Covid-19.*

1. Pendahuluan

Menurut laporan dari *We Are Social* pada bulan Januari 2022, terdapat peningkatan sebesar 12,35% dalam jumlah pengguna media sosial aktif di Indonesia dibandingkan dengan tahun sebelumnya,

yakni mencapai 191 juta orang. Berikut ini urutan media sosial mulai dari yang paling populer di Indonesia yaitu *Whatsapp, Instagram, Facebook, Tiktok, Twitter*, dan *Telegram*. Melihat begitu tingginya jumlah pengguna media sosial di Indonesia saat ini dapat dimanfaatkan untuk menganalisis sentimen opini publik di media sosial terhadap suatu topik / isu yang berkembang di masyarakat. Saat ini isu yang lagi hangat di media sosial yaitu "Perkuliahan *Hybrid*".

Pada akhir tahun 2019, muncul virus *Corona*, juga dikenal sebagai *Covid-19*, yang menyebar dengan cepat ke berbagai negara dan menjadi pandemi. Indonesia adalah salah satu negara yang terinfeksi *Covid-19*. Menurut situs *World Health Organization*, virus *Corona* adalah virus yang dapat menyebabkan penyakit pada manusia dan hewan. Pada manusia, virus *Corona* menyebabkan infeksi saluran pernapasan, mulai dari gejala flu biasa hingga *Middle East Respiratory Syndrome* (MERS) dan *Severe Acute Respiratory Syndrome*(SARS)[1].

Selama dua tahun terakhir, Indonesia telah dilanda pandemi *Covid-19* yang memberikan dampak buruk pada banyak sektor, termasuk sektor pendidikan. Penanganan yang rumit terhadap pandemi *Covid-19* menyebabkan pemerintah Indonesia menerapkan kebijakan-kebijakan yang sangat ketat untuk memutus mata rantai penyebaran virus tersebut. Kebijakan *Social Distancing* dan *Physical Distancing* adalah salah satu kebijakan yang memiliki dampak negatif terutama pada sektor pendidikan. Sejak *Covid-19* diumumkan sebagai bencana nasional, Menteri Pendidikan Nasional Indonesia mengeluarkan kebijakan untuk menutup sementara sekolah dan perguruan tinggi. Salah satu alternatif yang diterapkan oleh perguruan tinggi untuk tetap melaksanakan perkuliahan adalah dengan mengikuti kebijakan-kebijakan pemerintah, seperti melaksanakan perkuliahan jarak jauh secara *online*. [1].

Melalui perkuliahan jarak jauh secara *online*, dosen dan mahasiswa dapat melakukan kegiatan belajar mengajar tanpa harus bertemu secara langsung, menggunakan bantuan teknologi seperti *Zoom, Ms Teams, Webex*, dan sejenisnya. Meski begitu, dalam pelaksanaan perkuliahan online terdapat banyak kendala. Keterbatasan akses internet dan kurang terbiasanya mahasiswa maupun dosen menjadi salah satu kendala dalam perkuliahan *online*. Banyak keluhan baik dari mahasiswa ataupun orang tua mahasiswa terkait dengan pelaksanaan perkuliahan *online* [2].

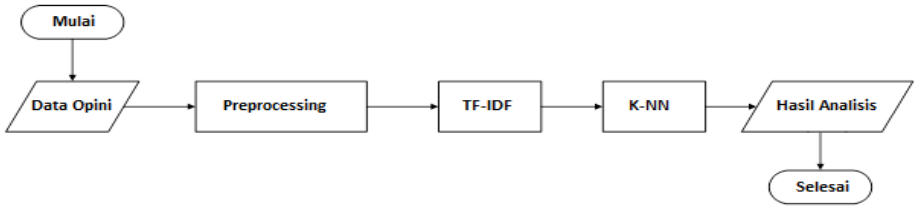
Seiring dengan menurunkan kasus *Covid-19* di Indonesia banyak Perguruan Tinggi mencari alternatif baru dalam mengatasi permasalahan pada perkuliahan *online* yaitu dengan melaksanakan perkuliahan secara *Hybrid*. Perkuliahan *Hybrid* adalah perkuliahan yang mengkombinasikan perkuliahan secara tatap muka dengan perkuliahan *online*. Pelaksanaan perkuliahan *Hybrid* tentunya memiliki kekurangan dan kelebihan yang disebabkan metode pembelajaran model ini baru pertama kali diterapkan. Perguruan Tinggi yang melaksanakan perkuliahan *Hybrid* ini perlu melakukan evaluasi terhadap metode pembelajaran tersebut guna untuk memperbaiki kekurangan-kekurangan yang ada. Salah satu cara mengevaluasi pelaksanaan Perkuliahan *Hybrid* dapat dilakukan dengan melakukan analisis sentimen opini publik terkait dengan Perkuliahan *Hybrid* [2].

Untuk melakukan analisis sentimen opini masyarakat yang terdapat di media sosial, dapat menggunakan bantuan *text mining*. Pada *text mining* memungkinkan dilakukan pengelompokan polaritas suatu *text* atau opini. Adapun tujuan dari pengelompokan tersebut yaitu untuk mengetahui polaritas suatu *text* atau opini apakah besifat positif, netral atau negatif. Penelitian tentang analisis sentimen sudah pernah dilakukan sebelumnya. Penelitian yang dilakukan S. Samsir dkk menemukan bahwa 69% sentimen negatif, 30% sentimen positif dan 1% sentimen netral tehadap pembelajaran daring. [3]. Penelitian yang dilakukan R. Kurniawan dan A. & Apriliani bertujuan untuk mengetahui polaritas publik terhadap virus corona [4]. Dari penelitian terkait yang telah dijabarkan hanya menggunakan satu media sosial yaitu twitter dan menggunaka algoritma *Naive Bayes* serta belum ada yang secara khusus meneliti terkait topik Perkuliahan *Hybrid*. Dalam penelitian yang lakukan, terdapat beberapa perbedaan dalam hal data opini yang terkait dengan Perkuliahan *Hybrid*, dan data yang diambil dari berbagai media sosial seperti *Twitter, Facebook*, dan *Instagram*. Algoritma *TF-IDF* dan *K Nearest Neighbor* digunakan dalam penelitian ini untuk memaksimalkan hasil analisis sentimen.

2. Metode Penelitian

Perancangan sistem yang dibuat merupakan sebuah perancangan untuk menganalisis sentimen opini publik pada sosial media terhadap Perkuliahan *Hybrid* dengan tujuan untuk mengetahui polaritas opini publik terhadap Perkuliahan *Hybrid*. Tahap awal akan dimulai dengan mengumpulkan data opini publik pada sosial media *Twitter*, *Facebook* dan *Instagram*. Adapun teknik yang digunakan dalam pengumpulan data yaitu *web scraper*. *Web Scraping* adalah proses ekstraksi informasi dari situs web dengan cara otomatis menggunakan perangkat lunak khusus yang disebut *scraper* atau *crawler*. Tujuan utama dari *Web Scraping* adalah untuk mengambil data yang terdapat pada suatu situs web dan

menyimpannya ke dalam format yang mudah dipahami oleh manusia atau komputer. Selanjutnya data tersebut melalui tahap *preprocessing*. Pada Tahap *Preprocessing* dilakukan beberapa tahapan diantaranya *case folding, remove punctuation, clean number, clean one character, convert word, remove stop word, tokenizing, stemming.* Kemudian, dilakukan implementasi algoritma *TF-IDF* untuk menghitung bobot dari setiap kata yang paling umum digunakan dalam *information retrieval*. [5]. Selanjutnya dilakukan implementasi *K-Nearest Neighbor* untuk mengklasifikasikan opini masyarakat terkait Perkuliahan *Hybrid* berdasarkan atribut dan *training sample*. Gambar 1 menujukan alur penelitian yang dilakukan [6].



Gambar 1 Metode Penelitian

2.1. Algoritma TF-IDF

Algoritma TF-IDF adalah metode untuk menentukan kepentingan suatu kata dalam suatu dokumen atau kumpulan dokumen. Algoritma ini menghitung nilai *Term Frequency* (TF) dan *Inverse Document Frequency* (IDF) dari suatu kata. TF menghitung seberapa sering kata tersebut muncul dalam dokumen, sedangkan IDF menghitung seberapa jarang kata tersebut muncul dalam seluruh kumpulan dokumen. Dengan menggabungkan nilai TF dan IDF, dapat dihitung nilai TF-IDF suatu kata dalam dokumen. Nilai TF-IDF yang tinggi menunjukkan bahwa kata tersebut penting dalam dokumen tersebut dan sering muncul dalam dokumen, tetapi tidak terlalu sering dalam seluruh kumpulan dokumen. Algoritma TF-IDF digunakan dalam pengolahan bahasa alami untuk mengekstrak kata kunci yang paling relevan dari dokumen. Rumus TF-IDF (*Term Frequency-Inverse Document Frequency*) adalah sebagai berikut:

$$TF - IDF\ (term, doc) = TF(term, doc) * IDF(term) \quad (1)$$

Keterangan :
- term adalah kata yang dihitung nilai TF-IDF-nya
- doc adalah dokumen yang dihitung nilai TF-IDF-nya
- TF(term, doc) adalah nilai Term Frequency dari kata tersebut di dalam dokumen
- IDF(term) adalah nilai Inverse Document Frequency dari kata tersebut dalam seluruh kumpulan dokumen

2.2. K-Nearest Neighbor

K-Nearest Neighbor (K-NN) merupakan salah satu algoritma supervised learning yang digunakan di dalam machine learning. Cara kerja algoritma K-NN adalah dengan menemukan tetangga terdekat dari data baru yang ingin diprediksi, berdasarkan data latih yang sudah ada sebelumnya. Dalam algoritma K-NN, setiap data dalam training data diwakili dalam sebuah ruang dimana setiap dimensi merepresentasikan sebuah fitur atau atribut dari data tersebut. Ketika ada data baru, algoritma akan mencari K data terdekat dalam training data dari data baru tersebut. Nilai K ditentukan sebelumnya. Untuk menemukan tetangga terdekat, umumnya algoritma K-NN menggunakan teknik perhitungan jarak Euclidean atau Manhattan. Jarak Euclidean menghitung jarak antara dua titik dalam ruang n-dimensi, sedangkan jarak Manhattan menghitung jumlah selisih absolut antara setiap dimensi dari dua titik.

$$d(p,q) = \sqrt{((p1 - q1)^2 + ((p2 - q2)^2 + \cdots + ((pn - qn)^2} \quad (2)$$

Keterangan :
- d(p,q) adalah jarak Euclidean antara titik p dan q
- p1, p2, ..., pn adalah nilai-nilai pada masing-masing dimensi dari titik p
- q1, q2, ..., qn adalah nilai-nilai pada masing-masing dimensi dari titik q

3. Hasil dan Pembahasan

3.1. Data Opini

Tabel 1 merupakan kumpulan data opini yang bersumber dari sosial media *twitter*, *facebook* dan *instagram*. Data opini yang digunakan dalam eksperimen ini sebanyak 20 data dari 200 data. Pada data opini ini masih terdapat berbagai spesial karakter, simbol dan kata yang tidak baku. Sebelum melakukan proses ke tahap selanjutnya data opini ini akan melalui proses preprocessing [7].

Tabel 1. Data Opini

Kode	Akun	Opini
D1	@itbfess	Katanya bakal hybrid nih kuliah, internet di kampus bagus ga sih? itb!
D2	@cornerhunter_	Jangankan anak SD yg baru masuk, gw pas kuliah hybrid k kampus sendirian, berasa bnget belajarnya, akhirnya jadi suka bnget matkul itu
D3	@haniepaboz	Semoga semester depan kuliah masih hybrid
D4	@dwldr25	Kuliah udh mulai hybrid semester depan. Gimana caranya aku bisa nyambi kerjasaran
D5	@wtrmel00n_	Tolong berbaik hati pada diriku yah 😔🙏 mental gue masih mental kuliah online soalnya 😔😔 ini bisa kuliah online lagi aja gak sih □□ atau enggak hybrid deh gapapa, jadi gak semua offline gitu □□
D6	@eleanorisk_	Seperti kelas pada umumnya. Pada semester pertama kelasku full online, sedangkan pada semester kedua sudah berjalan hybrid. Akan ada dokter yang mengajar/memaparkan materi disertai dengan slide PPT, diakhiri dengan sesi tanya jawab.
D7	@mocmiluvyu	Enakan 100% online, soalnya kalo hybrid harus kejer"an waktu gtu kadang masih di kampus tp ada kuliah yg online. Atau habis kuliah yg online ada yg offline dn wktunya kdg suka mepet gtu degdegan di jalan takut telat😔😔
D8	@Bilqisnabilaaa	Teruntuk rector dan dekan tercinte, yuu hybrid yuu, bosen laa online muluu
D9	@8123rio	Semoga besok2 mukjizat dating yaitu berupa pembatalan perkuliahan hybrid semua
D10	@_rawi_	Kirain yg hybrid iru Cuma mobil. Lah ini sekarang ada perkuliahan hybrid juga. Eh..iya kan hybrid kan ya? Etdah
D11	@tarooomess_	Tolong inimah plis banget buat em tolong ajukan tidaksetujuan saya atas rencana perkuliahan hybrid pasca uts dengan alasan nanggung banget dank lo mau balik pas libur lebaran pun kemungkinan susah dan tiketya sisa yg eksekutif jadi mahal dan saya tidak punya anggaran untuk tiket
D12	@ceavinrufus	Dari kuliah hybrid saya belajar bahwa setengah setengah itu gak baik
D13	@egar29_	Pendapatmu kuliah hybrid pie lur. Kok berat
D14	@neodoyies	kuliah hybrid ngang ngong ngang ngong, mending kalo offline sekalian 100%, yg kuliah dirumah ga denger dosennya ngomong apa 😔
D15	@pribadi_ihsan	Kuliah hybrid harapan saya : kuliahnya tatap muka, quiz & ujian online di rumah 😌
D16	@ninunenai	kuliah hybrid pertama yang sangat sangat tidak kondusif
D17	@UPIfess	!UPI Ya Allah 1am au kuliah hybrid, pengen kuliah online ajaaa 😔😔😔😔
D18	@dinsmelisa	Bisa ga ntar kalo udah normal kuliah hybrid aja, kuliah dan tutorial tetap pake zoom terus tramped dan praktikum aja yg luring
D19	@_nazdaa	Jujur kayaknya kalo kuliah hybrid aku blm siap hmm
D20	@purumpim	Iseng nanya ibu , gimana kalo semester depan kuliah hybrid tapi butuh surat izin dari ortu? "Gausah, kuliah disini aja udah" :(

3.2. Preprocessing

Tahap *preprocessing* merupakan suatu tahap dimana data opini melalui serangkaian proses yang bertujuan untuk mentransformasikan data mentah untuk siap dianalisis supaya dapat menghasilkan data yang akurat dan berkualitas. Pada eksperimen ini dilakukan beberapa tahap *preprocessing* diantaranya

1. *Case folding*: proses mengubah semua huruf dalam teks menjadi huruf kecil atau huruf besar
2. *Remove Punctuation*: proses menghapus tanda baca seperti titik, koma, tanda tanya, tanda seru, dan sebagainya dari teks.
3. *Clean number*: proses menghapus atau mengganti semua angka dalam teks dengan kata tertentu, seperti "number".

4. *Clean one character*: proses menghapus karakter tunggal atau simbol-simbol tertentu yang tidak berguna dalam teks.
5. *Convert word*: proses mengubah kata-kata dalam teks menjadi bentuk tertentu, seperti kata dasar, kata baku, atau kata dengan format tertentu.
6. *Remove stop word*: proses menghapus kata-kata umum atau *stop word* dalam teks.
7. *Tokenizing*: Proses memisahkan teks menjadi token atau unit-unit kecil, seperti kata-kata, frasa, atau simbol tertentu.
8. *Stemming:* proses mengubah kata-kata dalam teks menjadi bentuk dasarnya dengan menghapus imbuhan atau akhiran kata.

Pada Tabel 2 menunjukkan hasil dari tahapan preprocessing yang sudah dilakukan.

Tabel 2. Hasil *preprocessing*

Kode	Twitt	Label
D1	hybrid kuliah internet kampus bagus tidak itb	NETRAL
D2	anak sd kuliah hybrid kampus ajar suka matakuliah	POSITIF
D3	semoga semester depan kuliah hybrid	POSITIF
D4	kuliah hybrid semester depan sambil kerja saran	NETRAL
D5	tolong hati mental mental kuliah online soal kuliah online tidak tidak hybrid tidak apa-apa tidak semua offline	POSITIF
D6	kelas semester pertama kelas online penuh semester jalan hybrid dokter ajar papar materi slide ppt sesi tanya jawab	NETRAL
D7	enak online soal hybrid kejar kadang kampus kuliah online habis kuliah online offline kadang suka mepet degdeg jalan takut telat	NEGATIF
D8	rektor dekan cinta hybrid bosan online	POSITIF
D9	semoga besok mukjizat rupa batal kuliah hybrid	NEGATIF
D10	kira hybrid cuma mobil kuliah hybrid hybrid	NETRAL
D11	tolong mohon tolong aju tidak setuju rencana kuliah hybrid pasca uts alasan tanggung kamu balik libur lebar susah tiket sisa eksekutif mahal tidak punya anggaran tiket	NEGATIF
D12	kuliah hybrid ajar tengah tengah tidak baik	NEGATIF
D13	pendapat kamu kuliah hybrid berat	NEGATIF
D14	kuliah hybrid lebih baik offline kuliah rumah tidak dengar dosen bicara	NEGATIF
D15	kuliah hybrid harap kuliah tatap muka quiz ujian online rumah	POSITIF
D16	kuliah hybrid tidak kondusif	NEGATIF
D17	allah tidak mau kuliah hybrid kuliah online	NEGATIF
D18	bisa tidak normal kuliah hybrid kuliah tutorial tetap pakai zoom tramped praktikum luring	POSITIF
D19	jujur kuliah hybrid aku belum siap	NEGATIF
D20	iseng tanya ibu semester depan kuliah hybrid surat izin orangtua tidak perlu kuliah saja	NEGATIF

3.3. Pembobotan Kata

Pembobotan kata pada text mining adalah suatu proses di mana nilai atau bobot diberikan pada setiap kata dalam dokumen atau koleksi dokumen, dengan tujuan untuk mengidentifikasi kata-kata yang paling relevan dan penting dalam dokumen. Dengan demikian, pembobotan kata membantu dalam analisis teks dan informasi, dan memudahkan pengambilan kesimpulan dan pemahaman terhadap isi dokumen tersebut. Metode TF-IDF digunakan untuk menghitung bobot dari setiap kata yang paling umum digunakan. Setelah melalui tahapan preprocessing, hasil sampel akan dijadikan kamus kata-kata yang menjadi acuan untuk perhitungan nilai bobot. Tabel 3 menunjukkan hasil pembobotan kata-kata tersebut.

Tabel 3. Pembobotan Kata

	tf											df	D/df	IDF	IDF+1
	D1	D2	D3	D4	D5	D6	D7	D8	D9	...	D20				
hybrid	1	1	1	1	1	1	1	1	1	...	1	22	0,9	-0,045	0,955
kuliah	1	1	1	1	2	0	2	0	1	...	2	24	0,8	-0,096	0,904
internet	1	0	0	0	0	0	0	0	0	...	0	1	20	1,301	2,301
kampus	1	1	0	0	0	0	1	0	0	...	0	3	6,7	0,826	1,826
bagus	1	0	0	0	0	0	0	0	0	...	0	1	20	1,301	2,301
suka	0	1	0	0	0	0	1	0	0	...	0	2	10	1	2
semoga	0	0	1	0	0	0	0	0	1	...	0	2	10	1	2
online	0	0	0	0	2	1	3	1	0	...	0	9	2,2	0,342	1,342
offline	0	0	0	0	1	0	1	0	0	...	0	3	6,7	0,826	1,826
kejar	0	0	0	0	0	0	1	0	0	...	0	1	20	1,301	2,301
takut	0	0	0	0	0	0	1	0	0	...	0	1	20	1,301	2,301
bosan	0	0	0	0	0	0	0	1	0	...	0	1	20	1,301	2,301
mukjizat	0	0	0	0	0	0	0	0	1	...	0	1	20	1,301	2,301
batal	0	0	0	0	0	0	0	0	1	...	0	1	20	1,301	2,301
setuju	0	0	0	0	0	0	0	0	0	...	0	1	20	1,301	2,301
susah	0	0	0	0	0	0	0	0	0	...	0	1	20	1,301	2,301
mahal	0	0	0	0	0	0	0	0	0	...	0	1	20	1,301	2,301
berat	0	0	0	0	0	0	0	0	0	...	0	1	20	1,301	2,301
kondusif	0	0	0	0	0	0	0	0	0	...	0	1	20	1,301	2,301

3.4. Hasil Perhitungan TF-IDF

Setelah dilakukan pembobotan kata-kata tersebut, kemudian dihitung nilai TF-IDF. Tabel 4 menunjukkan hasil perhitungan nilai TF-IDF.

Tabel 4. Nilai TF-IDF

D1	D2	D3	D4	D5	D6	D7	D8	D9	D10	D11	D12	...	D20
0,95	0,95	0,95	0,95	0,95	0,95	0,95	0,95	0,95	2,86	0,95	0,95	...	0,95
0,9	0,9	0,9	0,9	1,81	0	1,81	0	0,9	0,9	0,9	0,9	...	1,81
2,3	0	0	0	0	0	0	0	0	0	0	0	...	0
1,83	1,83	0	0	0	0	1,83	0	0	0	0	0	...	0
2,3	0	0	0	0	0	0	0	0	0	0	0	...	0
0	2	0	0	0	0	2	0	0	0	0	0	...	0
0	0	2	0	0	0	0	0	2	0	0	0	...	0
0	0	0	0	2,68	1,34	4,03	1,34	0	0	0	0	...	0
0	0	0	0	1,83	0	1,83	0	0	0	0	0	...	0
0	0	0	0	0	0	2,3	0	0	0	0	0	...	0
0	0	0	0	0	0	2,3	0	0	0	0	0	...	0
0	0	0	0	0	0	0	2,3	0	0	0	0	...	0
0	0	0	0	0	0	0	0	2,3	0	0	0	...	0
0	0	0	0	0	0	0	0	2,3	0	0	0	...	0
0	0	0	0	0	0	0	0	0	0	2,3	0	...	0
0	0	0	0	0	0	0	0	0	0	2,3	0	...	0
0	0	0	0	0	0	0	0	0	0	2,3	0	...	0
0	0	0	0	0	0	0	0	0	0	0	0	...	0
0	0	0	0	0	0	0	0	0	0	0	0	...	0

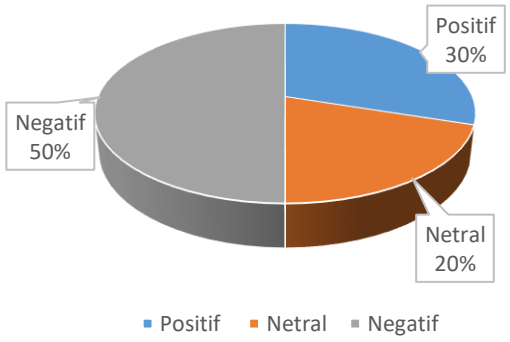
3.5. Implementasi K-Nearest Neighbor

Setelah nilai *TF-IDF* untuk setiap kata dalam dokumen diperoleh, langkah selanjutnya adalah menghitung jarak *Euclidean Distance* antara dokumen berdasarkan nilai-nilai *TF-IDF* yang ada di Tabel

Nilai *TF-IDF*. Perhitungan jarak dimulai dari kata pertama hingga kata terakhir. Untuk menghitung jarak, nilai *TF-IDF* suatu kata dikurangi dengan nilai *TF-IDF* kata lainnya, kemudian hasilnya dikuadratkan. Selanjutnya, semua nilai kuadrat tersebut dijumlahkan dan akar dari jumlah tersebut dihitung. Berikut ini adalah contoh perhitungan jarak antar kata.

$$\textit{Euclidean distance}\ (D1\text{-}D2) = \sqrt{(0,9-0,9)^2} + \sqrt{(0.9-0.9)^2} + \sqrt{(2,3-0)^2} + \sqrt{(1,8-1,8)^2} \quad (3)$$
$$= 0 + 0 + 5,29 + 0$$
$$= 5,29$$

Gambar 2. Menunjukan pie chart menunjukkan hasil klasifikasi sentimen dari 200 data uji, dengan 20% data diberi label netral (laporan), 50% diberi label negatif (keluhan), dan 30% diberi label positif (dukungan).



Gambar 2. Hasil Analisis Sentimen Perkuliahan Hybrid

4. Kesimpulan

Dari penelitian yang telah dilakukan menggunakan text mining untuk menganalisis sentimen media sosial terhadap Perkuliahan Hybrid, dapat disimpulkan bahwa terdapat sentimen positif sebesar 30%, netral sebesar 20%, dan negatif sebesar 50% terhadap Perkuliahan Hybrid.

Daftar Pustaka

[1] Siahaan dan Matdio, "Dampak Pandemi Covid-19 Terhadap Dunia Pendidikan," *Jurnal Kajian Ilmiah*, vol. 20, no. 2, pp. 1-3, 2020.

[2] Nafrin, I. Aulia dan H. Hudaidah, "Perkembangan pendidikan Indonesia di masa pandemi COVID-19," *Edukatif: Jurnal Ilmu Pendidikan*, vol. 3, no. 2, pp. 456-462, 2021.

[3] S. Samsir, A. Ambiyar, U. Verawardina, F. Edi dan R. Watrianthos, "Analisis Sentimen Pembelajaran Daring Pada Twitter di Masa Pandemi COVID-19 Menggunakan Metode Naïve Bayes," *Jurnal Media Informatika Budidarma*, vol. 5, no. 1, pp. 157-163, 2021.

[4] R. Kurniawan dan A. & Apriliani, "Analisis sentimen masyarakat terhadap virus corona berdasarkan opini dari Twitter berbasis web scraper," *Jurnal INSTEK (Informatika Sains dan Teknologi)*, vol. 5, no. 1, pp. 67-75, 2020.

[5] Ma'arif dan A. Aziz, "Penerapan Algoritma Tf-Idf Untuk Pencarian Karya Ilmiah," *Jurnal. Jurusan Teknik Informatika. Fakultas Ilmu Komputer. Universitas Dian Nuswantoro Semarang*, vol. 5, no. 4, 2015.

[6] R. I. Ndaumanu, "Analisis Prediksi Tingkat Pengunduran Diri Mahasiswa dengan Metode K-Nearest Neighbor," *Jatisi*, vol. 1, no. 3, 2014.

[7] L. a. I. M. Hermawan, "Pembelajaran text preprocessing berbasis simulator untuk mata kuliah information retrieval," *Jurnal Transformatika*, vol. 17, no. 2, pp. 188-199, 2020.