https://jurnal-umbuton.ac.id/index.php/taksonomi
https://doi.org/10.35326/taksonomi.v4i1.6962

Taksonomi: Journal of Basic Education Research
Volume 4 Number 2 Year 2024

E - ISSN: 2798-947X
P - ISSN: 2986-6499

The Use of Machine Learning to Evaluate the Jatim Cerdas and Sehat Program: A Case Study in East Java

Aries Dwi Indriyanti[1]*, Rahmat Gernowo[2], Eko Sediyono[3]

[1,2,3]Universitas Diponegoro, Indonesia

*email Correspondence: ariesdwi@students.undip.ac.id

Abstracts

This study evaluates the effectiveness of the Jatim Cerdas and Sehat Program using a quantitative approach based on machine learning to deliver quick, accurate, and data-driven insights. By combining sentiment analysis with Geographic Information Systems (GIS), the research uncovers public opinions, key issues affecting program implementation, and the spatial distribution of education and healthcare problems in East Java. The study employs a quantitative methodology, collecting data from Twitter through data crawling via the Twitter API. Sentiment analysis is conducted using the Naive Bayes algorithm, while spatial analysis utilizing GIS is employed to visualize the geographic spread of the issues. The sentiment analysis findings reveal that public perception of the program is largely negative. Key concerns identified include the unfairness of the education zoning system, confusing curriculum changes, high education costs despite promises of free education, and ongoing debates over the National Examination. In healthcare, complaints focus on disparities in healthcare facilities, inconsistent service delivery, and high treatment costs, including issues related to BPJS. Spatial analysis highlights priority areas, particularly rural and suburban regions, that require further attention regarding access to education and healthcare.

Keywords: Machine Learning, Sentiment Analysis, Geographic Information Systems

1. Introduction

The Jatim Cerdas dan Sehat Program is a strategic initiative by the East Java Provincial Government aimed at improving the quality of life of the community by enhancing access to quality education and adequate healthcare services (https://kominfo.jatimprov.go.id/berita). The program was launched to reduce social inequality, improve human resource quality, and support sustainable development. Education and healthcare are two crucial sectors in human development (Arifin, 2015; Nurlita, 2017; Herdiansyah, 2020). A good quality education will produce productive and innovative human resources, while access to adequate healthcare services will improve the quality of life and economic productivity. Siregar (2024) argues that quality education can produce competent, ethical human resources capable of competing in the era of globalization. Education also helps learners develop attitudes, skills, and intellectual abilities. Improving the quality of education and healthcare is a primary priority in regional development. Equal access to education and healthcare for all citizens is a crucial element in the success of regional development (Mochtar et al., 2014; UNESCO, 2015; Sachs, 2015).

The Jatim Cerdas dan Sehat Program is a policy initiative by the East Java Provincial Government designed with an inclusive approach, particularly targeting low-income communities, to achieve equitable access to public services. East Java, as one of the largest provinces in Indonesia with a population of over 40 million, faces significant development challenges due to its diversity in demographics, geography, and socio-economic conditions (Central Statistics Agency, 2021). Inequality in access to education and healthcare has become a major issue, especially in underdeveloped rural and urban areas. Therefore, the launch of this program is expected to serve as one of the solutions to achieving the nation's goals of equitable education and healthcare for all citizens. The Jatim Cerdas dan Sehat Program consists of two main pillars: Tis-Tas (Free and Quality) in the education sector, and Desa Sehat (Healthy Villages) in the healthcare sector. In the education sector, the program aims to provide equitable and quality education access, particularly for low-income communities, through policies such as tuition fee waivers, improving the quality of educators, and supporting school accreditation. On the other hand, the healthcare sector focuses on equitable healthcare services by developing healthcare facilities in remote areas, improving access to medical professionals, and providing nutrition education to the public (Dindik Jatim, 2021; Detik.com, 2023).

Although the Jatim Cerdas dan Sehat Program aims to increase inclusive access to education and healthcare, its implementation still faces various challenges. Public dissatisfaction with the execution of this policy often surfaces on social media platforms, such as Twitter, which has become a primary channel for expressing public opinion directly and openly. Data from social media presents a significant opportunity for the government to better evaluate the program. Social media enables the government to capture public opinions without intermediaries, including criticism, complaints, or recommendations related to program implementation (Sauter & Pereira, 2015; Chadwick & Dennis, 2019). The data obtained from social media can be analyzed in real-time, allowing the government to respond quickly to emerging issues, such as access disparities or service-related problems. In today's digital age, the use of technologies like machine learning becomes a highly effective tool for governments to evaluate programs and policies (Bins, 2019; Kharbanda, 2020; Salsabila, 2024). In this context, sentiment analysis and GIS can be particularly useful. Sentiment analysis and trending topics on social media can help the government identify specific issues, such as protests against zoning, education costs, or complaints about healthcare services. Meanwhile, social media data combined with Geographic Information Systems (GIS) can be used to map areas frequently reporting particular issues, such as a lack of schools or healthcare services. As a result, the government can formulate data-driven policies to allocate resources more effectively.

This research is highly significant as it provides a rapid and accurate data-driven evaluation. This approach allows the government to evaluate programs dynamically, rather than relying solely on bureaucratic reports that are often slow or incomplete. With data-driven evaluation, the government can respond more quickly and accurately to societal changes and needs (Gonzales, 2018; Aslam, 2023). One of the technologies used in this study is sentiment analysis and spatial data analysis using Geographic Information Systems (GIS). Sentiment analysis offers insight into public satisfaction with government programs. By collecting data from various sources, such as Twitter social media, sentiment analysis provides a clearer picture of public opinion regarding

the policies being implemented. This helps the government identify critical issues faced by the community and provides valuable input for making adjustments or improvements to existing policies. In this way, the government can enhance its responsiveness to emerging issues and improve public trust in the policies being applied. Spatial data on the distribution of issues can also be geographically analyzed, helping the government identify priority areas requiring quicker intervention. By utilizing big data and analytical technologies, the government can design more effective policies that align with the real needs and conditions of society. Therefore, the research questions are: 1) What is the public perception of the effectiveness of the Jatim Cerdas dan Sehat Program? 2) What are the main issues affecting the program's implementation in urban and rural areas? 3) What is the geographical distribution of educational and health issues in East Java based on GIS analysis?

Several previous studies, such as the one by Nova (2023) titled "Comparison of the K-Nearest Neighbor (K-NN) and Decision Tree Methods in Sentiment Analysis of JNE Expedition Application Reviews on Google Play Store," demonstrate how internet technology has transformed the business landscape by driving the growth of e-commerce. The findings suggest that the Decision Tree model outperforms KNN in classifying user reviews related to JNE's expedition service. Overall, this research offers insights into how classification methods can be used to analyze customer reviews and understand sentiment related to expedition services. Another study by Aditya et al. (2016) examined bot spammer applications and legitimate users by integrating sentiment analysis based on emotions and time interval entropy. The combined knowledge-based and machine learning-based approach was used to classify tweets with positive, negative, or neutral sentiment. The results showed that the precision and recall of the proposed method reached 83% and 91%, respectively, proving that combining Sentiment Analysis (SA) and Time Interval Entropy (TIE) can optimize system performance in detecting bot spammers.

Kusumaningrum et al. (2014) also explored an unsupervised machine learning method, Latent Dirichlet Allocation (LDA) in Integrated Visual Vocabulary (IVV). The study found that using IVV improved overall accuracy (OA) by 11-12% and 6% when applied to selected features and all features, respectively. The selected features, such as CIELab color moments and GLCM, provided better OA than using CIELab color moments or GLCM individually. Additionally, the results showed that the OA of LDA exceeded that of C4.5 and Naive Bayes trees by about 20%, indicating that this machine learning method can effectively map language. Previous research contributed to the effectiveness of unsupervised machine learning in processing language data, which is crucial for sentiment analysis in public policy. However, this study enhances the data by incorporating spatial data analysis through the application of GIS to evaluate government policy programs. The sentiment analysis and GIS-based approach is a relatively new and innovative method in public policy evaluation. Therefore, the findings of this research could serve as a model or reference for other local governments to evaluate development programs using modern technology.

2. Research methods

The research approach used is quantitative, employing machine learning algorithms. The quantitative approach allows us to analyze data in the form of numbers and information, as well as identify trends in sentiment changes over time (Bryman, 2016; Creswell, 2018; Kearney, 2019). By using quantitative data, we can also compare

the level of public satisfaction with various policies. For example, we can compare the public's satisfaction with education policies and healthcare policies. The results of quantitative research can be generalized to a broader population, providing a more accurate representation of public opinion.

During this stage, data collection involves gathering posts, comments, and tweets from social media platforms. The online data will be compiled into a dataset, with each news article treated as an individual document. The subsequent step is sentiment classification, aimed at finding the most accurate classification model. After the data undergoes text preprocessing and feature selection, it is split into training and testing sets using the 10-fold cross-validation technique. This method ensures that the training and testing occur ten times, with data divided into ten segments. Finally, a GIS application is employed to map the geographic origins of the tweets, enabling the identification of issues based on the regional map.

After collecting the data, the next step involves analyzing it using different machine learning techniques to uncover relevant patterns and connections. This phase includes: 1) Data Preprocessing: The data collected from surveys, social media, and documents is processed to remove noise and ensure it is ready for analysis. This involves text cleaning, deduplication, and categorizing the data; 2) Sentiment Analysis: Machine learning algorithms such as Naive Bayes or Support Vector Machine (SVM) are used to analyze social media data and classify public sentiment (positive, negative, or neutral) towards the Jatim Cerdas dan Sehat program; 3) Issue Classification: Algorithms like K-Means clustering and LDA are applied to identify the main issues affecting the implementation of the program, based on public opinion data; 4) Geospatial Analysis: GIS technology combined with machine learning will be used to analyze the distribution of educational and healthcare issues in East Java and identify regions that require more attention.

The results of the analysis will be presented in an easily understandable format for policymakers and the public. Data presentation will include data visualization, such as tables and geospatial maps, to illustrate sentiment analysis, issue classification, and the distribution of educational and healthcare problems. Additionally, the findings will be detailed in a narrative report that outlines the key discoveries from the study, including policy recommendations based on the analysis results.

3. Results and Discussion

3.1 Public Perception and Key Issues of the Jatim Cerdas dan Sehat Program

Based on the research findings, public perception of the effectiveness of the Jatim Cerdas dan Sehat Program, as derived from sentiment analysis, remains low. Many tweets express dissatisfaction with the program's effectiveness, particularly in the areas of education and healthcare. The data collection process for the Jatim Cerdas Sehat program began with crawling data from Twitter using the API. The search keywords used, such as "Jatim cerdas sehat," "zonasi," "fasilitas pendidikan," "pelayanan kesehatan," "rumah sakit," and "puskesmas," were aimed at capturing public opinion about the policy. From the crawling process, a total of 1,965 tweet data points were collected. However, after removing duplicate data, the number of usable tweets dropped to 1,045. This data was then processed through a preprocessing phase to ensure its quality before being used in the data mining process.

Table 1. Crawling Data Results for Jatim Cerdas dan Sehat

id_str	username	created_at	full_text
1854108247202119761	MYF69	Fri Dec 13 10:25:08 +0000 2024	Lebih baik ditiadakan aja. Percuma ada program tsb kalo sumber dananya dr kenaikan pajak rakyat. Lbh baik pendidikan gratis. Bisa meringankan pengeluaran ortu/wali sehingga malah bisa fokus bekerja untuk memperbaiki gizi mereka sendiri. Dananya bisa buat seragam & buku2.
1853348152285544860	Eradotid	Wed Nov 06 17:27:47 +0000 2024	[NEWS EXPLAINED] Kabar bagus? Mendikdasmen Abdul Mu'ti rencanakan untuk mengembalikan Ujian Nasional (UN) tingkat SD & SMP di tahun ajaran baru. Perihal sistem zonasi UN, serta UN dan kurikulum merdeka ia ungkap masih dikaji apakah mereka perlu dihapus atau tidak. https://t.co/Sw8gjdecRH
1853396504943432136	yuhlookeasy	Mon Nov 04 11:18:55 +0000 2024	kalo gue setuju tetap hapus un dan hilangkan/kurangi sistem zonasi jadi masuk smp ato sma lewat hasil nilai rapot yg sidanira itu agar lebih mudah meningan jatah jalur prestasi dan jalur zonasi dituker
1853343551977091145	injeunithread	Mon Nov 04 07:48:30 +0000 2024	canza @elcanzasfr ·Jul 10, 2021 to @cinpus_ uang aku belum cukup buat beli buku yang mahal. Biaya pendidikan sma aku juga mahal kak karena ada biaya spp yang cukup mahal/bulan sekitar 150k aku ada di dom kediri dan papa aku kerja harus pp dari surabaya kediri setiap hari biaya transport juga nggak murah kak-
1855073086888136819	pakeMotorTrail	Sat Nov 09 02:21:04 +0000 2024	karena ada biaya spp yang cukup mahal/bulan sekitar 150k aku ada di dom kediri dan papa aku kerja harus pp dari surabaya kediri setiap hari biaya transport juga mahal kak apalagi pandemiDengan gaji yang cuma 2-3 juta cuma cukup buat biaya transport,pendidikan,makan
1854065693366603871	booblebee_	Wed Nov 06 07:38:02 +0000 2024	@ferrykoto @ganjarpranowo and 2 others bernasib lebih miris,pendatang(dr wilayah) yg bekerja di Surabaya. Anak2nya wajib sekolah swasta. Dimana tdk

			ada SPP GRATIS. Uang Pendaftaran Mahal Daftar Ulang biaya selanggit
1855025658625507497	football_garin	Sun Nov 10 14:07:42 +0000 2024	Banyak anak Indonesia yang masih mengalami stunting. Gizi buruk ini harus segera diatasi! #stunting #giziburuk
1855613305534210090	widodol_joko	Thu Nov 07 14:35:04 +0000 2024	Program pemerintah untuk mengatasi masalah gizi harus efektif. #programgizi #kesehatan
1854533028426772892	tresnaUmmuHanan	Sun Dec 08 18:58:14 +0000 2024	Sebelumnya, pemerintah sudah meluncurkan program makan siang bergizi gratis (MBG) bagi anak-anak dan ibu hamil. Program yang disebut-sebut akan menurunkan kasus stunting ini anggarannya kini diciutkan dari Rp15 ribu menjadi Rp10 ribu per orang.
1854346592578592875	deuxyribose	Sun Nov 10 22:34:12 +0000 2024	vibes rumah sakit non bpjs tuh kenapa berkelas bgt ya sepi suster ramah2 mau lu request atau nanya hal bodou pun mereka tetep jawab dengan senyum coba rumah sakit bpjs dr satpamnya kek tai susternya jg senyum gapernah hih. https://t.co/T2Ca8CAbkv
1855740773213650968	Afflictionist	Fri Nov 08 09:06:44 +0000 2024	Kenapa semua perokok malah dicover BPJS nebulnya ya semakin pd buat hidup panjang semakin naikin dagu
1854873437447315897	gadogadopare	Thu Nov 07 04:37:02 +0000 2024	Keuntungan orang menengah ke atas pakai pelayanan kesehatan BPJS yaitu: 1. Bisa dirawat di ruang VIP hanya tambah selisih bayar kamar saja sudah bisa nikmati fasilitas VIP 2. Pengobatan bisa lebih cepat karena diagnosis diketahui lebih cepat Mereka biasanya begitu sakit langsung
1854382528091587048	helmieh_	Mon Nov 11 04:35:36 +0000 2024	Pentingnya sistem pelayanan kesehatan yang efisien dan terjangkau tidak bisa dipandang sebelah mata. Beberapa negara telah berhasil menciptakan model pelayanan kesehatan yang tidak hanya memenuhi kebutuhan masyarakat, tetapi juga menjamin kualitas dan aksesibilitas.

Table 2. Stemming Results for Jatim Cerdas (Left) and Jatim Sehat (Right)

	stem_text
0	[lebih, baik, tiada, aja, percuma, program, ts...
1	[maklum, ibu, megawati, enggan, makan, makan, ...
2	[news, explained, kabar, bagus, mendikdasmen, ...
3	[kalo, gue, tuju, tetap, hapus, un, hilangkank...
4	[un, pelan, baik, kualitas, soal, tuju, asesme...
5	[mendikdasmen, abdul, muti, sedang, timbang, k...
6	[canza, jul, to, uang, aku, cukup, buat, beli,...
7	[biaya, spp, cukup, mahalbulan, k, aku, dom, d...
8	[and, others, nasib, lebih, mirispendatangdr, ...
9	[elhakim, , mar, kunjung, bang, sandi, pantai,...

	stem_text
0	[banyak, anak, indonesia, alami, stunting, giz...
1	[program, perintah, atas, masalah, gizi, efektif]
2	[belum, perintah, luncur, program, makan, sian...
3	[vibes, rumah, sakit, non, bpjs, tuh, kelas, b...
4	[kenapa, semua, okok, malah, dicover, bpjs, ne...
5	[untung, orang, tengah, atas, pakai, layan, se...
6	[penting, sistem, layan, sehat, efisien, jangk...
7	[gue, tuju, sama, dapat, harus, semua, orang, ...
8	[korban, woohwan, jatuh, bangsal, penuh, gak, ...
9	[dengan, dukung, infrastruktur, baik, distribu...

From the Twitter data table above, it can be observed that topics such as nutritious meals, the National Examination (UN), and zoning were the most discussed in the education sector, while issues like stunting, malnutrition, and healthcare facilities were the trending topics in the health sector. The issue of free lunch dominated the conversation, as its implementation has not been carried out effectively, leading the public to express negative sentiments. Education zoning was also widely discussed, particularly regarding the restriction of high-achieving students' access to preferred schools. This criticism mainly came from urban areas with high population density. Frequent curriculum changes also attracted attention from netizens, with many complaints from the public expressing confusion over sudden changes. These tweets highlight how instability in education policy can hinder the teaching and learning process. Teachers struggle to adjust teaching methods, while students have insufficient time to adapt to the new curriculum. Many tweets also reflect that the public still faces high educational costs, particularly to access quality education. Private schools with adequate facilities and qualified teachers are often the preferred choice, but their fees are unaffordable for many families. The debate surrounding the National Examination (UN) is still prevalent on Twitter, where some support its removal due to the psychological burden it imposes on students, while others believe the UN is essential for measuring educational quality standards. These tweets primarily come from students and parents sharing their experiences on Twitter. From this data, it can be concluded that education-related issues in East Java have not been fully addressed with the launch of the Jatim Cerdas program. Therefore, the government needs to consider several steps, such as improving the quality of schools across all zones, maximizing the nutritious meal program, providing incentives for schools with low student numbers, and ensuring that zoning policies are more flexible to accommodate the individual needs of students and families. By implementing these steps, the Jatim Cerdas program could be more effective in improving the quality of education without sacrificing the principles of equity and fairness.

Table 3. Data Labeling Results for Jatim Cerdas (Left) and Jatim Sehat (Right)

Stemming	Score	Klasifikasi
Dana bantuan	1	Positif
Zonasi	-1	Negative
Biaya mahal	-1	Negative
Makan siang Gratis	-1	Negative
Gaji naik	1	Positif
Fasilitas Kesehatan	-1	Negative

Before constructing the data mining model, an initial analysis was carried out using a word cloud to identify the most commonly mentioned terms in discussions about the program. Words like "zoning," "lunch," "high costs," "health facilities," and "UN" were prominent, highlighting the key concerns of the public. For instance, "Zoning" points to concerns over the controversial educational zoning system, "lunch" pertains to the government's school meal program for children, and "health facilities" reflects discussions on the availability and quality of healthcare services. From the data labeling results, it was found that 66.28% of the text was classified as negative, and 33.72% as positive. To better understand public sentiment prior to developing the data mining model, the data will be visualized in a word cloud to display the most frequently discussed words related to the Jatim Cerdas Sehat program and their frequency of occurrence. Figure 4.6 illustrates that the words most often mentioned in relation to the Jatim Cerdas Sehat program are "zoning," "lunch," "high costs," "health facilities," and "UN." These terms frequently appeared in tweets as key issues in public conversations. While the sentiment analysis from Twitter revealed some positive tweets about the program, the majority of the tweets reflected dissatisfaction and a perception of the program's failure among the East Java community.

3.2 Geographic Distribution of Education and Health Issues in East Java Based on GIS Analysis

The geographic distribution of education and health issues in East Java can be analyzed using spatial data by leveraging Geographic Information Systems (GIS) technology. By collecting data on education and healthcare accessibility, such as the number of schools and health facilities per district, as well as graduation rates or disease prevalence in each area, this study can provide a more detailed depiction of existing problems. This data can then be mapped to highlight areas with more significant issues that require special attention. Based on the research findings, issues related to zoning and healthcare facilities, in particular, received negative responses in geographically remote and economically weaker regions. For instance, areas like Bangkalan, Sampang, and Probolinggo are often marginalized when it comes to access to quality education and healthcare. In the context of the Jatim Cerdas program, regions such as Surabaya, Malang, Sidoarjo, and Madiun emerged as areas showing improvements in educational performance. This is evident from various tweets on Twitter mentioning that students in these areas often achieve outstanding academic accomplishments. This improvement can be attributed to several key factors, such as the presence of top-tier private schools with adequate facilities, access to quality teachers, and better access to educational technology. However, even though these regions have succeeded in improving educational performance, they also face significant controversy regarding the school zoning policy. Zoning, which aims to place

students in schools closer to their homes and reduce the gap between schools, has received strong criticism in these cities.

Figure 1. Wordcloud

From the Twitter data table above, it can be seen that education zoning is the most frequently discussed topic. Many tweets mention public dissatisfaction with this policy, with the majority of criticisms coming from urban areas with high population density. Complaints are most prevalent in regions like Sidoarjo and Malang, where the gap between public and private schools has become more apparent.

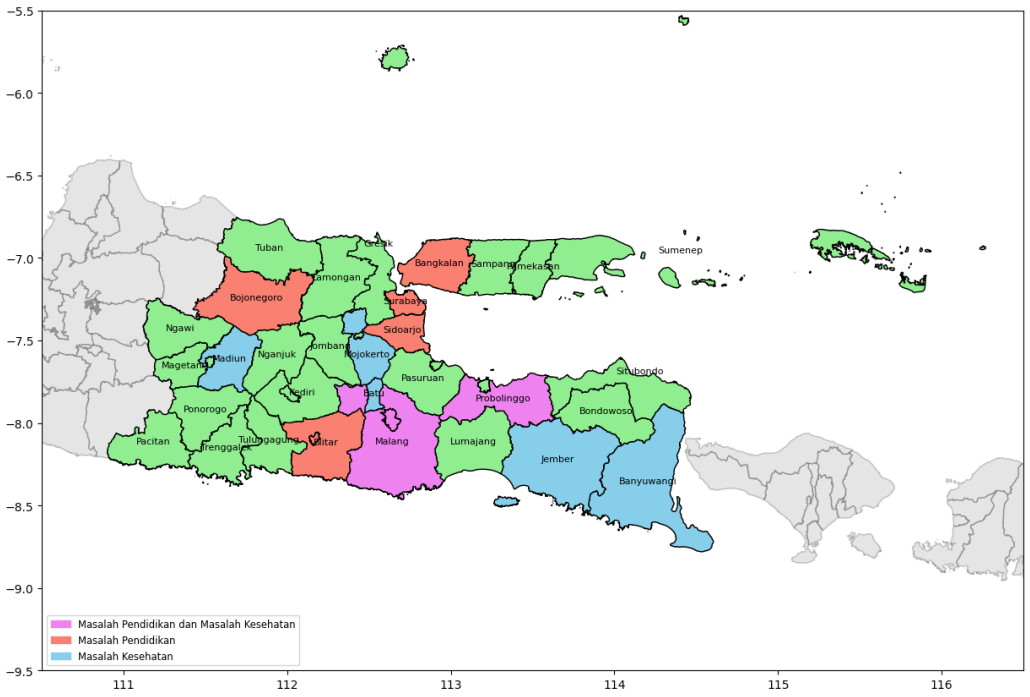

Figure 2. GIS Map of Jatim Cerdas Sehat

Based on the GIS analysis and public sentiment, the distribution of educational and healthcare issues in East Java can be visualized through regional mapping. In major cities, tweet data indicates that the main focus is on the issues of school zoning and the high costs of education at private schools. Meanwhile, in coastal and rural areas such as Bangkalan, Bojonegoro, Probolinggo, and Blitar, public tweets discuss disparities in access to education. Equity in education is the main concern, with complaints about the lack of school facilities, such as classrooms, laboratories, libraries, and the shortage of qualified teachers. Additionally, supporting infrastructure, such as access roads to schools and internet connectivity, presents significant barriers to the teaching and

learning process. The zoning issue in areas like Blitar, Bangkalan, Bojonegoro, and Probolinggo remains a concern, but its impact is not as severe as in urban areas. According to GIS data and sentiment analysis, zoning is a central issue in larger cities. Many tweets highlight that high-achieving students are often unable to attend preferred schools due to zoning restrictions. This causes frustration among parents who feel that the system prioritizes administrative rules over individual merit. Many tweets, especially from Surabaya and Malang, point out that the cost of education at private schools, which are perceived to offer better quality, is a heavy burden for middle-class families. Parents feel forced to choose private schools because they perceive public schools as unable to meet educational expectations. This situation exacerbates the educational gap between wealthy and less affluent families, with children from lower-income households often relegated to lower-quality schools, while wealthier children have access to premium educational facilities.

In contrast, tweets from coastal and rural areas focus more on the inequality in access to education. Education equity is a primary concern, with many people expressing frustration over the lack of opportunities for their children to access quality schools. Schools in these areas often lack basic facilities, such as adequate classrooms, laboratories, libraries, and qualified teachers. Limited infrastructure is a significant barrier to teaching and learning. Issues such as difficult access to schools, shortages of learning materials, and a lack of supporting facilities like internet access are frequently mentioned in tweets. In Probolinggo and Bangkalan, for instance, many students must travel long distances to reach schools due to the lack of nearby institutions. Meanwhile, Blitar and some areas of Sidoarjo complain about inadequate school infrastructure. The zoning issue in Blitar, Bangkalan, Bojonegoro, and Probolinggo is also a concern, but it is less significant compared to larger cities. Zoning is considered less relevant in these areas due to the absence of high-quality alternative schools within the same zone. With this data mapping, the GIS visualization, combined with public sentiment data, helps to identify priority regions. This allows for more targeted policy decisions to address educational issues in East Java, both in urban and rural areas.

In the context of Jatim Sehat, positive sentiment is largely dominated by urban areas such as Surabaya, Malang, Madiun, Mojokerto, and Batu. Sentiment analysis based on Twitter posts shows positive responses related to healthcare facilities, the low incidence of stunting, and the availability of specialist medical personnel. Many tweets express satisfaction with the hospitals and healthcare services, as well as with consistent nutrition education programs. The presence of quality hospitals and adequate medical staff in these areas has received positive feedback from the public. GIS visualization reveals a concentration of adequate healthcare facilities in these regions. The map displays the distribution of hospitals, Puskesmas (community health centers), and Posyandu (integrated health service posts), ensuring that the public can easily access healthcare services. This aligns with the GIS mapping and sentiment analysis, where tweets with positive tones predominantly come from these regions. Well-developed infrastructure has enhanced public perception. Areas like Mojokerto, with the lowest supplementary feeding programs and nutrition education campaigns. The Innovative Health Program in Batu, a tourist area, has also garnered many positive tweets, highlighting the integration of health initiatives with environmental wellness programs, such as eco-friendly healthcare facilities and community-based tourism.

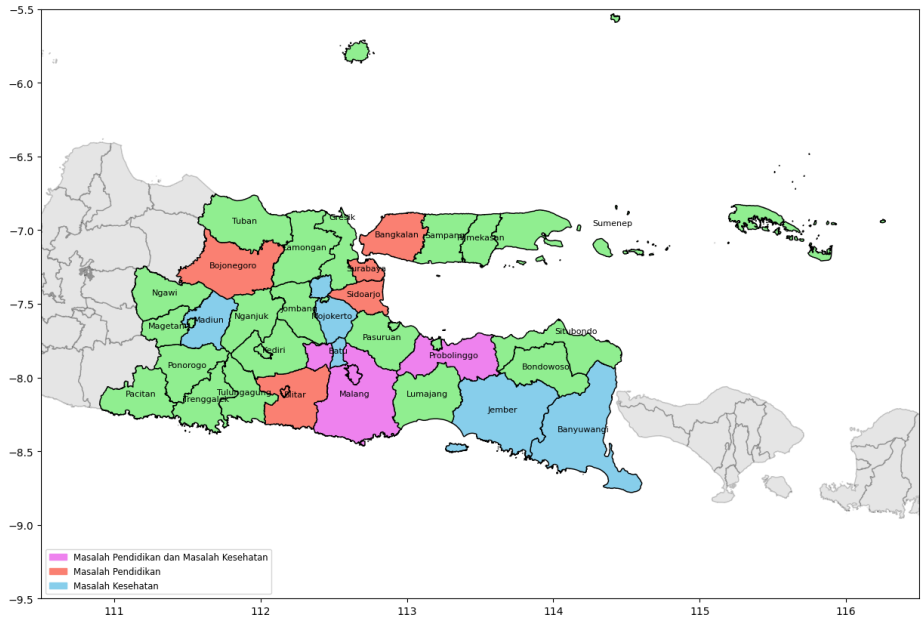

Figure 3. GIS Map of Jatim Cerdas Sehat

The areas with negative responses to the Jatim Sehat program highlight significant challenges, particularly in rural and coastal regions such as Jember, Banyuwangi, Madura, and Probolinggo. In Jember and Banyuwangi, the main issues include the high prevalence of stunting, which is caused by a lack of nutrition education, limited access to healthy food, and disparities in healthcare service distribution. Furthermore, the geographical factor of large rural areas with difficult access to healthcare facilities further hampers the effectiveness of the program. Remote areas, including Madura, Jember, and Probolinggo, face major difficulties in accessing basic healthcare services, especially in villages far from facilities like Puskesmas and hospitals.

In areas such as Pamekasan and Probolinggo, the gap in healthcare services is a primary concern, with uneven distribution of facilities and a shortage of specialist doctors, such as at Waru Hospital and RS Mohammad Noer. This is compounded by low public awareness of the importance of nutrition and healthy living behaviors, which adds complexity to the health issues in these regions. Meanwhile, in Nganjuk, the high prevalence of stunting and limitedhealthcare infrastructure represent major challenges. The GIS map reveals an unequal distribution of healthcare facilities in rural areas, where the long distances between communities and Puskesmas or hospitals impede access to healthcare services. Sentiment data also shows public complaints about the lack of specialist doctors, especially in Madura regions like Pamekasan, Bangkalan, and Sumenep. The GIS map shows the concentration of specialist doctors in major cities such as Surabaya and Malang, with a significant gap in remote areas. This disparity in medical staff distribution creates negative perceptions, especially in areas heavily dependent on the Desa Sehat program and specialist doctor visits. The combination of GIS and sentiment data highlights that geographic factors and insufficient healthcare infrastructure are key obstacles in implementing health programs, necessitating special attention to service equity and public education improvements.

4. Conclusion

The use of machine learning in analyzing the Jatim Cerdas and Sehat Program has proven to provide significant contributions in evaluating the effectiveness of government programs quickly, accurately, and based on data. By combining sentiment analysis, issue classification, and spatial analysis based on Geographic Information Systems (GIS), this study was able to identify public perceptions, uncover key issues affecting program implementation, and map the geographical distribution of educational and health problems in East Java. Based on sentiment analysis, public perception of this program is still low, with more negative sentiments compared to positive ones. Several key issues that emerged include the education zoning system, which is perceived as unfair, where high-achieving students often fail to enter preferred schools due to zoning limitations. This is particularly felt in rural areas, where access to quality schools is limited. Additionally, frequent curriculum changes have also been a point of criticism. The frequent shifts in curriculum are seen as confusing for both students and teachers, hindering the stability of the learning process. Another issue frequently raised is the high cost of education, despite the program promoting free education. Many feel that costs remain high, especially for schools with good facilities and teachers, further exacerbating the inequality in access to education. The debate over the National Examination (UN) is also a contentious topic, with some supporting its removal, arguing that it burdens students, while others consider it an important standard for measuring the quality of education. Additionally, the perceived decline in the quality of public schools compared to decades ago has increased public concern, while the cost of private schools is considered unaffordable for many families. In terms of health, the primary complaints focus on the inequity in healthcare facilities, services, and the costs associated with medical treatment and BPJS (Indonesian health insurance). Overall, these complaints reflect the public's frustration with the effectiveness of the Jatim Cerdas and Sehat program's implementation. Many feel that the program has not fully addressed the challenges on the ground, whether in terms of equal access to education, quality improvement, or aligning policies with the needs of the community. Through GIS-based geographical mapping, this study has identified priority areas that require more immediate intervention, such as remote areas with limited access to education and healthcare facilities. This information is crucial for policymakers to design more focused programs that cater to the specific needs of each region.

References

Aditya, Christian Sri Kusuma., Mamluatul Hani'ah., Alif Akbar Fitrawan., Agus Zainal Arifin. (2016). Deteksi Bot Spammer pada Twitter Berbasis Sentiment Analysis dan Time Interval Entropy. *Jurnal Buana Informatika* 7(3)

Arifin, M. Y., Suswandi, P. E., & Adenan. (2015). Pengaruh pengeluaran pemerintah sektor kesehatan, pengeluaran pemerintah sektor pendidikan dan pertumbuhan ekonomi terhadap indeks pembangunan manusia Jawa Timur 2006-2013. *Universitas Jember*.

Aslam, F. (2023). The impact of artificial intelligence on chatbot technology: A study on the current advancements and leading innovations. *European Journal of Technology*, 7(3), 62–72. https://doi.org/10.47672/ejt.1561

Azmi, F. (2023, February 15). *4 tahun Khofifah-Emil, Muhammadiyah puji program Jatim Cerdas dan Sehat*. DetikJatim. https://www.detik.com/jatim/berita/d-6569249/4-tahun-khofifah-emil-muhammadiyah-puji-program-jatim-cerdas-dan-sehat

Binns, A., & Duggan, P. (2019). Machine learning and the role of artificial intelligence in government. *Journal of Public Administration Research and Theory*, 29(1), 65-81. https://doi.org/10.1093/jopart/muy024

Bryman, A. (2016). *Social research methods* (5th ed.). Oxford University Press.

Chadwick, A., & Dennis, J. (2019). Social media, political polarization, and the role of government. *Political Studies Review*, 17(1), 73-87. https://doi.org/10.1177/1478929918773896

Creswell, J. W. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dinas Pendidikan Provinsi Jawa Timur. (n.d.). *Badan Pendapatan dan Organisasi Pendidikan Provinsi Jawa Timur*. https://dindik.jatimprov.go.id/pages/bpopp

González-Pizarro, M., & Alvarado, C. (2018). The impact of machine learning algorithms on public policy decision-making. *Policy and Society*, 37(3), 393-413. https://doi.org/10.1080/14494035.2018.1473452

Herdiansyah, D., & Kurniati, P. S. (2020). Pembangunan sektor pendidikan sebagai penunjang indeks pembangunan manusia di Kota Bandung. *Jurnal Agregasi (Jurnal Aksi Reformasi Government Dalam Demokrasi)*, 8(1), Mei. https://doi.org/10.34010/agregasi.v8i1.2765

https://dindik.jatimprov.go.id/pages/bpopp

Indonesia Statistics Bureau (BPS). (2021). *Statistik sosial ekonomi: Profil demografi dan sosial ekonomi Jawa Timur*. Badan Pusat Statistik.

Kearney, M. W. (2019). *Quantitative research methods for social sciences*. SAGE Publications.

Kharbanda, S., & Singh, D. (2020). Using machine learning for policy evaluation: Insights from big data. *International Journal of Information Technology and Decision Making*, 19(6), 1717-1735. https://doi.org/10.1142/S0219622020500545

Kusuma, A. S., Hani'ah, M., Fitrawan, A. A., & Arifin, A. Z. (2016). Deteksi bot spammer pada Twitter berbasis sentiment analysis dan time interval entropy. *Jurnal Buana Informatika*, 7(3), 212-220. https://doi.org/10.24002/jbi.v7i3.656

Kusumaningrum, R., Wei, H., Manurung, R., & Murni, A. (2014). Integrated visual vocabulary in latent Dirichlet allocation-based scene classification for IKONOS image. *Journal of Applied Remote Sensing*, 8(1), 083690. https://doi.org/10.1117/1.JRS.8.083690

Mochtar, N. E., Hendarman, H., Indrastuti, N., Suryana, C., Restuningsih, K., & Tartila, S. L. (2014). *Pendidikan untuk pembangunan berkelanjutan (Sustainable Development) di Indonesia*. Komisi Nasional Indonesia untuk UNESCO (KNIU): Kementerian Pendidikan dan Kebudayaan. ISBN: 978-602-70365-1-2

Nova, Putri. (2023). Perbandingan metode K-Nearest Neighbor (K-NN) dan Decision Tree pada analisis sentimen ulasan penggunaan aplikasi expedisi JNE di Google Play Store. *Skripsi Universitas Pendidikan Indonesia*.

Nurlita, Pertiwi. (2017). *Implementasi sustainable development di Indonesia*. Pustaka Ramadhan.

Sachs, J. D. (2015). *The Age of Sustainable Development*. Columbia University Press.

Salsabila, Thalita., Indrawati, T. M., & Fitrie, R. A. (2024). Meningkatkan efisiensi pengambilan keputusan publik melalui kecerdasan buatan. *Journal of Internet and Software Engineering*, 1(2), 1-21.

Sauter, M., & Pereira, A. (2015). Social media and public opinion: New channels for citizen engagement. *International Journal of Public Administration*, 38(12), 901-911. https://doi.org/10.1080/01900692.2015.1077639

Siregar, R. V., Lubis, P. K. D., Azkiah, F., & Putri, A. (2024). Peran penting pendidikan dalam pembentukan sumber daya manusia cerdas di era digitalisasi menuju smart society 5.0. *IJEDR: Indonesian Journal of Education and Development Research*, 2(2), 15-27. https://doi.org/10.1234/ijedr.2024.2.2.15

United Nations Educational, Scientific and Cultural Organization (UNESCO). (2015). *Education for Sustainable Development: A Roadmap*. UNESCO.