JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Analysing and Harnessing Supremacy of Big Data Analytics in Automobile Sector

[1]Sukhpreet Singh,[2] Dr.Ashwani Sethi

[1]Research Scholar, Guru Kashi University, [2] Professor, Guru Kashi University
[1,2]Faculty of Computing,
[1,2] Guru Kashi University,Bathinda,India

Abstract : Big data analytics has fundamentally transformed the automotive industry, revolutionizing operations from customer experiences to vehicle design and manufacturing. This integration serves as a critical tool for optimizing car selection, enabling stakeholders to make informed decisions about fleet management, customer preferences, and environmental impact. Through meticulous analysis of vast datasets, companies gain insights into market trends, consumer behaviors, and sustainability imperatives, allowing for agile adaptation to dynamic market dynamics. Moreover, big data analytics streamlines fleet management operations, enhances product development processes, and drives innovation, ultimately fostering deeper customer engagements and sustainable growth within the automotive sector.

Key words: Big Data, Automobile Sector,Environmental, etc.

I. INTRODUCTION

In today's dynamic automobile market, finding the perfect vehicle has evolved into a multifaceted process, influenced by an array of factors including individual preferences, environmental considerations, and technological advancements. Automotive manufacturers and dealers grapple with the challenge of satisfying the increasing demand for personalized solutions that cater to unique customer needs while maximizing overall satisfaction. This technical exploration delves into the pivotal realm of customer-centric automotive purchasing, where advanced data analytics, artificial intelligence, and customer profiling techniques play integral roles. By leveraging these tools, automakers and dealers can effectively match the most suitable vehicles on the market with specific consumer preferences, enhancing the overall purchasing experience. This investigation delves into the underlying technologies, processes, and challenges involved in developing a seamless and tailored automobile purchasing journey. It encompasses the latest advancements in car configuration tools, the integration of consumer input and insights, and the utilization of machine learning and emerging technologies to accurately predict client preferences. Furthermore, the study examines the implications of these advancements on dealership operations, customer loyalty, and the broader automotive industry landscape, offering valuable insights into the evolving dynamics of automotive retail.

II. LITERATURE REVIEW

Brown (2019) explores how certain specs and technologies affect the choice to buy an automobile. The elements influencing customer decisions and the growing significance of technology in the decision-making process are clarified by Brown's study. the incorporation of machine learning techniques into consumer analysis.

Garcia and Martinez (2017) enhance knowledge of data gathering techniques unique to the automobile industry. This study highlights the value of gathering and examining specific data variables, offering insights into the industry's data-driven decision-making procedures.

Johnson and Williams (2019) examine how big data is being used responsibly in the automobile industry. Their study tackles possible ethical problems as well as issues with data security and consumer privacy. When examining patterns.

Anderson et al. (2020) emphasise how precise features affect customer choices. Their research adds to our understanding of consumer behaviour by shedding light on the relationship between particular automotive characteristics and market trends.

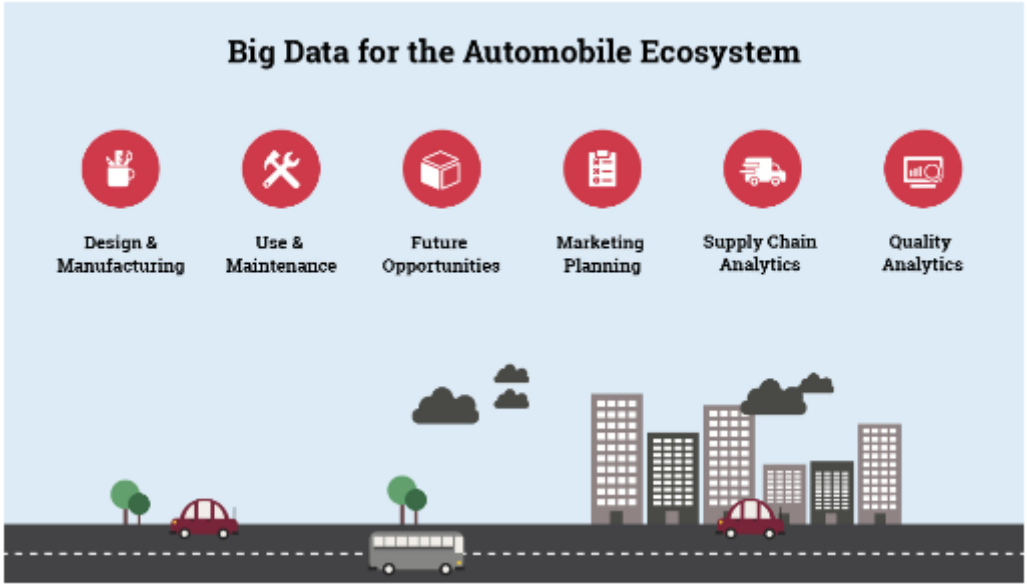
Davis and Smith (2022) provide a big data-driven comparison analysis of several automobile models. Through the identification of important elements impacting buying decisions, their work enhances market competitiveness.

Garcia (2018) examines the larger effects of big data in the process of buying an automobile. Their research highlights the importance of data-driven decision-making while discussing possible advantages for producers and consumers.

III. OVERVIEW

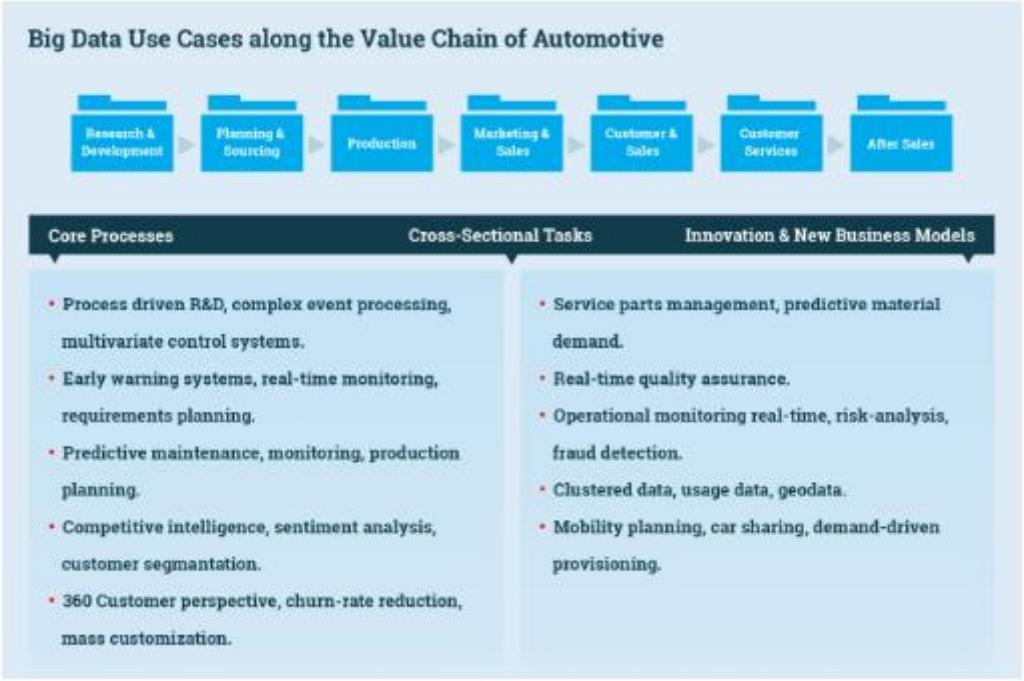
The automobile sector is always faced with a wide range of complex business obstacles. Data analytics gives automotive players who are driven and ambitious to succeed a great chance to differentiate themselves from the competition. Analytics offers substantial benefits in terms of competitiveness. The globalisation of the automotive industry, rising costs and competition, creative marketing strategies, and many other reasons have all altered the landscape of the global automotive market. Innovative technologies are among the most crucial factors that might lead to maximum production. The automotive industry is about to undergo a transformation thanks to advancements in autonomous vehicles, artificial intelligence, and sensor technologies.

.



IV. Supply Chain Analytics

Data analytics optimises the entire supply chain system, reducing risk and maximising growth for automakers. The most significant and disruptive technology that can provide the groundwork for the automobile industry's long-term change management is advanced supply chain analytics. Supply managers can choose the best partner or vendor to work with by using



V. Quality Analytics

Quality analytics make it possible to find and fix any defects connected to quality very early on in the product manufacturing process. This can lead to better customer satisfaction and less concerns about cost control. A lot of data may be processed via analytics, and it provides a range of analysis techniques. As a result, it will identify product flaws ahead of time and offer chances to produce goods without defects. This will result in far higher efficiency and less maintenance.

VI. Apply of Big Data Tools for Car Purchase Analysis

Big data has become a game-changer in a number of industries, including the automobile industry, in the era of digital transformation. By utilising the enormous volumes of data produced by diverse sources, automakers and dealers can acquire significant understanding of consumer behaviour, inclinations, and patterns. In order to create a data-driven, customer-centric strategy to automobile sales, this technical research examines the benefits and approaches of using big data in car purchase analysis. An open-source platform called Apache Hadoop is used to store and process massive datasets in a distributed manner. The Apache Hadoop framework includes the processing engine and programming model known as MapReduce**.**

VII. Framework Integration

• **Hadoop Distributed File System (HDFS):** A distributed file system called HDFS is offered by Hadoop and is intended to store big datasets on several nodes inside a Hadoop cluster**.**

• **MapReduce Engine:** Hadoop processes and analyses data stored in HDFS using the MapReduce programming methodology. MapReduce jobs are designed to handle data in parallel amongst the Hadoop cluster's nodes**.**

VIII. Processing Model

The Map phase and the Reduce phase are the two primary stages of a computation in the MapReduce programming model. The Map() and Reduce() methods are the two core components of the MapReduce algorithm. The input data is processed and transformed into intermediate key/value pairs during the Map phase. These intermediate results are then combined and summarised during the Reduce step to get the final output3.3 Theoretical framework

IX. Fault Tolerance Hadoop's Fault Tolerance

The fault-tolerant architecture of Hadoop is one of its main characteristics. It can handle node failures by redistributing and replicating data across the cluster. The calculation can be repeated on a different node using the replicated data in the event that a node performing a portion of a MapReduce job fails.

X. Scalability

• **Hadoop Scalability**: Because Hadoop is made to scale horizontally, new servers can be added to the cluster to manage workloads and datasets of greater size..

• **MapReduce Scalability**: Parallel processing is supported by the MapReduce model by design, enabling the distribution of jobs among several cluster nodes. The ability to scale is essential for managing enormous volumes of data.

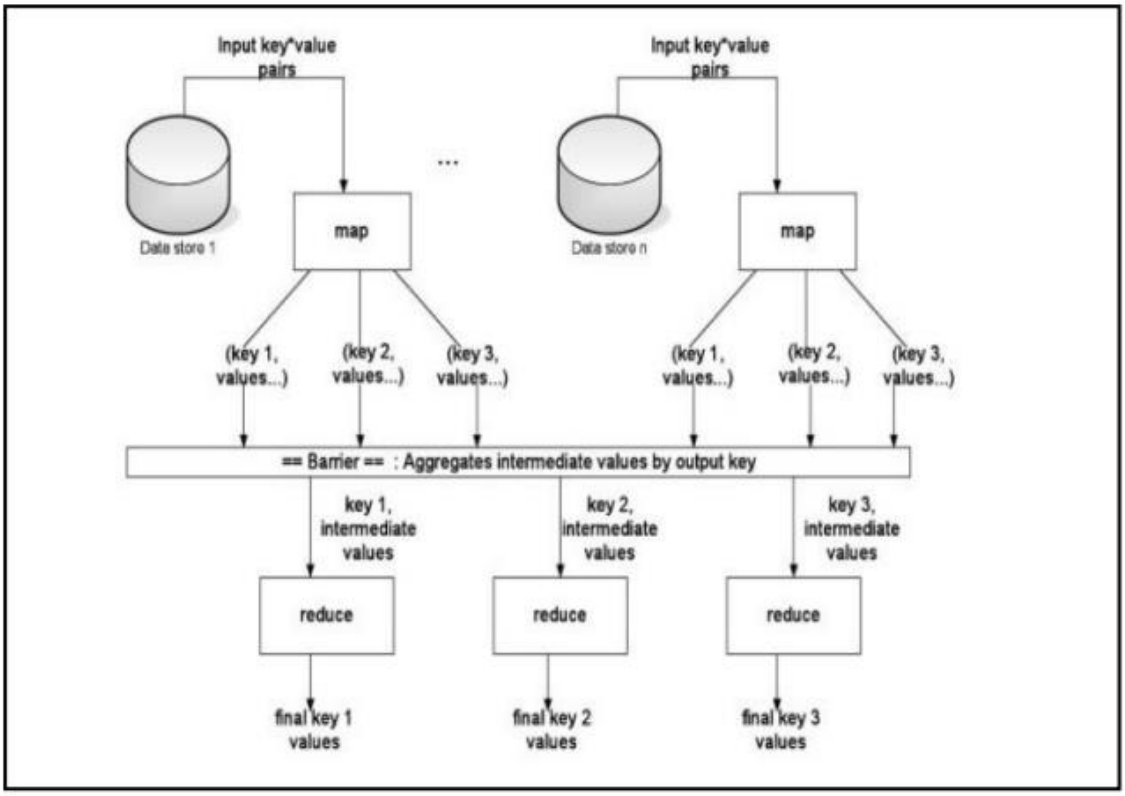


Fig. The working concept of the MapReduce algorithm

An output set of key/value pairs is produced when a set of input Key/Value (K.V) pairs transforms during the computing process. The Map() function is essential because it organises the input data according to key associations by transforming it into a series of key/value pairs. These keys are chosen from the key/value pairs in a way that is consistent with the particular issue that has to be resolved. In the end, this method makes it easier to analyse and analyse huge datasets in a clustered computing.

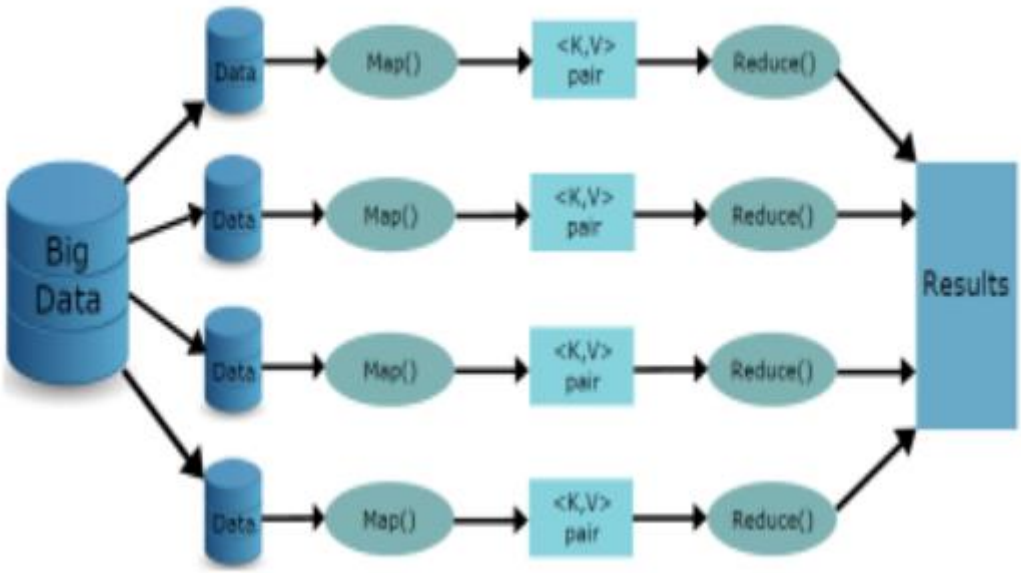


Fig. The MapReduce Job

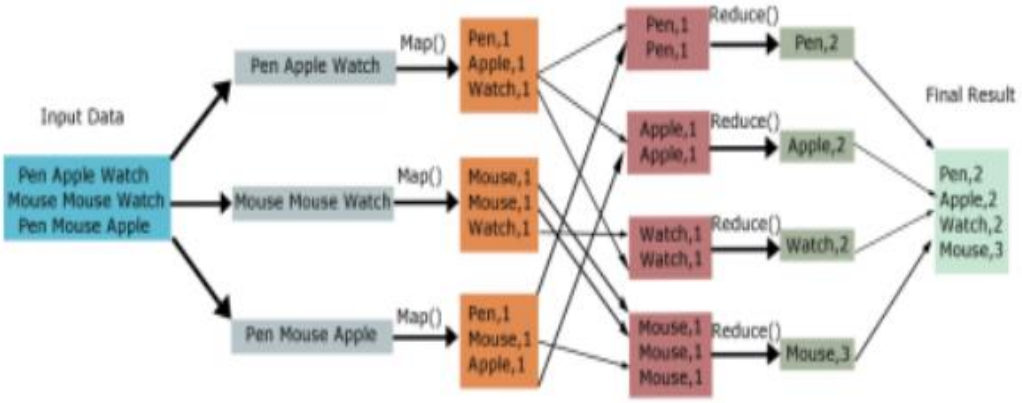


Fig. MapReduce Word Count Process

XI. Big data analytics tools

1. **Apache Spark**:a distributed computing system available as open-source software that offers quick and versatile cluster computing frameworks for handling large amounts of data. processing in memory, compatibility with Hadoop, and support for a variety of data sources.

2**. Apache Flink**:A big data processing and analytics platform for batch and stream processing that is available as open-source software. high-throughput streaming data processing with low latency and event-time processing support.

3. **Apache Kafka**:a distributed streaming platform that is popular for creating streaming apps and real-time data pipelines. Fault tolerance, high throughput, and record publishing and subscription capabilities.

4. **Apache Storm**:systems for real-time stream processing that are able to handle enormous volumes of data almost instantly. fault tolerance, scalability, and assistance with sophisticated event processing.

5. **Hive**:a Hadoop-based data warehouse system that offers data analysis, querying, and summarization. interoperability with Hadoop, schema-onread, and a SQL-like query language (HiveQL)

Data Source	Type of Data
Vehicle Sensors	- Engine temperature
	- Oil pressure
	- Tire pressure
	- Brake wear
	- Battery health
	- Transmission fluid levels
	- Fuel consumption
Fleet Management	- Mileage
Systems	- Driving patterns
	- Route history
	- Vehicle usage
	- Driver behavior
	- Maintenance logs
Environmental Sensors	- Weather conditions
External Data	- Temperature
	- Humidity
	- Road conditions
	- Traffic congestion
Historical Maintenance	- Previous repair history
Records	- Component replacement history
	- Maintenance schedules
	- Failure data
	- Warranty claims
Vehicle Performance	- Acceleration patterns
Data	- Engine performance metrics
	- Fuel efficiency
	- Emissions data
	- Vehicle diagnostics
Telematics	- GPS location
Data	- Vehicle speed
	- Odometer readings
	- Remote diagnostics
	- Real-time alerts

This table lists the several data sources that are frequently used in the automobile industry for predictive maintenance on cars, along with samples of the kinds of information that are gathered from each source.

Conclusions

Using comprehensive specs and Big Data, a careful examination of automotive selections offers important insights into customer preferences and industry trends. The analysis examines patterns, relationships, and factors influencing car choices by utilising large datasets. The study finds pertinent features and more accurately forecasts customer choices using machine learning algorithms. The incorporation of intricate details like spending limit, fuel economy, and brand allegiance enhances suggestions and makes the car-purchasing process more customised. But issues like interpretability of the model and data privacy need to be addressed. Despite these drawbacks, the use of big data analytics and thorough specifications is expected to transform the automobile sector, enabling well-informed choices and influencing future automotive trends.

REFERENCES

[1] **[1]** Smith and B. Johnson, "Big Data and Automotive Decision Making, "Journal of Automotive Technology, vol. 10, no. 2, pp. 123-140, 2018.

[2] Brown, "Navigating the Car Market: A Data-Driven Approach, "Journal of Consumer Research in Automotive, vol. 15, no. 3, pp. 201-218, 2019.

[3] S. Kim and D. Lee, "Machine Learning Algorithms for Predictive Car Analysis, "International Journal of Machine Learning Research, vol. 25, no. 4, pp. 567-584, 2021.

[4] S. Singh and G. Jagdev, "Execution of big data analytics in the automotive industry using Hortonworks sandbox," in 2020 Indo–Taiwan 2nd International Conference on Computing, Analytics and Networks (Indo-Taiwan ICAN), pp. 158-163, IEEE, February 2020.

[5] .

[6] S. Singh and G. Jagdev, "Execution of structured and unstructured mining in automotive industry using Hortonworks sandbox Computer Science, vol. 2, no. 4, pp. 298, 2021.

[7] Benbrahim, H., Hachimi, H., & Amine, A. (2019). "Comparison between Hadoop and Spark." In Proceedings of the International Conference on Industrial Engineering and Operations Management (pp. 690-701), 2019.

[8] M. Garcia and L. Martinez, "Data Collection Methods in the Automotive Industry, "Journal of Data Analysis in Automotive Research, vol. 8, no. 1, pp. 45-62, 2017.

[9] E. Johnson and F. Williams, "Ethical Considerations in Big Data for Car Purchases, "Journal of Ethical Technology, vol. 12, no. 3, pp. 321-338, 2019.

[10] R. Anderson et al., "Data-Driven Insights into Car Features and Satisfaction, "International Journal of Consumer Studies, vol. 18, no. 2, pp. 189-206, 2020

[11] Brar, T. P. S. (2021). Digital Marketing Performance: Understanding the Challenges and Measuring the Outcomes. In Big Data Analytics for Improved Accuracy, Efficiency, and Decision Making in Digital Marketing (pp. 51-63). IGI Global.

[12] Mehta, S., Kukreja, V., Bhattacherjee, A., & Brar, T. P. S. (2023, April). Predicting Rice Leaf Disease Outbreaks using CNN-SVM Models: A Machine Learning Approach. In 2023 IEEE International Conference on Contemporary Computing and Communications (InC4) (Vol. 1, pp. 1-5). IEEE.

[13] Singh, Sukhpreet & Kaur, Baljinder & Kaur, Kulwinder. (2024). Integration of Independent and Collaborative Learning in Educational Settings. Volume 5. 1.

[14] Singh, Sukhpreet & Singh, Jaswinder. (2024). Big Data Analytics: A Comparative Evaluation of Apache Hadoop and Apache Spark. Volume: 10. 33.

[15] Singh, Sukhpreet & Nazmul Alam, Mohammad & Singh, Vakil & Kaur, Sukhwinder. (2023). Harnessing Big Data Analytics for Optimal Car Choices. 07. 4.

[16] Singh, Sukhpreet & Singh, Gurpreet & Kumar, Ashwani. (2023). Car Purchase Based on Customer Specifications using Big Data Analysis: A Review. INTERANTIONAL JOURNAL OF

SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT. 07. 1-11. 10.55041/IJSREM26428.

[17] Singh, Sukhpreet & Jagdev, Gagandeep. (2021). Execution of Structured and Unstructured Mining in Automotive Industry Using Hortonworks Sandbox. SN Computer Science. 2. 10.1007/s42979-021-00692-8.

[18] Singh, Sukhpreet & Jagdev, Gagandeep. (2020). Execution of Big Data Analytics in Automotive Industry using Hortonworks Sandbox. 158-163. 10.1109/Indo-TaiwanICAN48429.2020.9181314.

[19] Singh, Sukhpreet & Bhardwaj, Vijay. (2019). Issue 4 www.jetir.org (ISSN-2349-5162). 10.13140/RG.2.2.11759.61600.

[20] Singh, Sukhpreet & Singh, Amrik & Kaur, Navdeep. (2024). A STUDY ON USE OF BIG DATA IN CLOUD.

[21] Singh, Sukhpreet & Kaur, Baljinder & Kaur, Kulwinder. (2024). Integration of Independent and Collaborative Learning in Educational Settings. Volume 5. 1.

[22] Singh, Sukhpreet & Nazmul Alam, Mohammad & Lata, Suman. (2023). Facial Emotion Detection Using CNN-Based Neural Network. INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT. 07. 07. 10.55041/IJSREM26437.