



COLOMBIA
POTENCIA DE LA
VIDA

UNIT 3

THE WONDERS OF THE MODERN TECHNOLOGY



TIC

► TALENTO
TECH

Región 3 Cauca Y Nariño

INTELIGENCIA
ARTIFICIAL

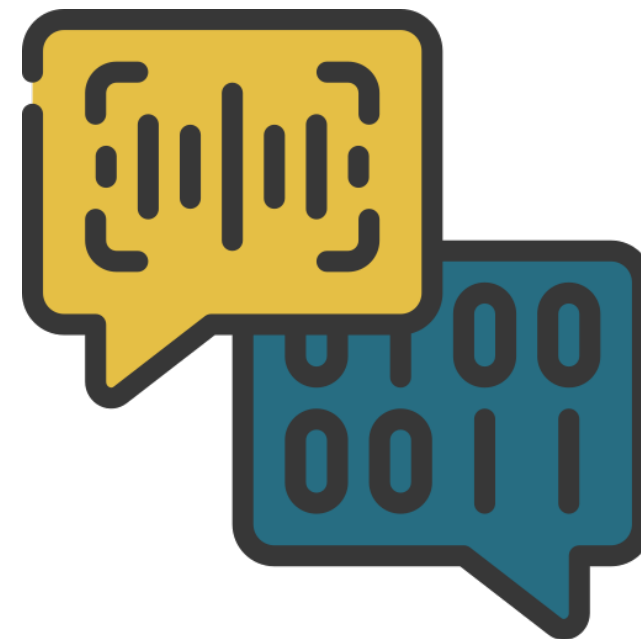
JAVA

EDUCACIÓN



OBJECTIVES OF THE UNIT

- Develop precise and effective language skills for articulating data-related concepts and insights as well as Acquire the ability to apply appropriate language conventions in various data communication contexts.
- Develop proficiency in the language of Natural Language Processing (NLP), encompassing the ability to understand and use key terms, techniques and concepts effectively within the AI context.



BAMBOOZLE

AI REVIEW

<https://www.baamboozle.com/game/634234>



HAVE A BUG

Refers to encountering a glitch or problem in a computer program or system.

*"After the software update, users reported that the application would frequently freeze—looks like we **have a bug** to fix in the code."*

WHAT IS DATA ANALYTICS?

Data analytics is the science of analyzing raw data to make conclusions about information. Many of the techniques and processes of data analytics have been automated into mechanical processes and algorithms that work over raw data for human consumption.



USEFUL DATA ANALYTICS TERMINOLOGY

DATA ANALYSTS

They are data professionals who gather, clean, study, or interpret data in order to solve business problems. They tend to work alongside other data analytics professionals, such as data scientists and data engineers.

BIG DATA

Refers to the huge amounts of data that large businesses and other organizations collect and store

NETWORK ANALYTICS

The science of describing and, especially, visualizing the connections among objects. The objects might be human, biological or physical. Graphical representation is a crucial part of the process.

MACHINE LEARNING

Analytics in which computers “learn” from data to produce models or rules that apply to those data and to other similar data.

Data Analytics: What It Is, How It's Used, and 4 Basic Techniques

KEY TAKEAWAYS

- Data analytics is the science of analyzing raw data to make conclusions about that information.
- Data analytics help a business optimize its performance, perform more efficiently, maximize profit, or make more strategically-guided decisions.
- The techniques and processes of data analytics have been automated into mechanical processes and algorithms that work over raw data for human consumption.
- Various approaches to data analytics include looking at what happened (descriptive analytics), why something happened (diagnostic analytics), what is going to happen (predictive analytics), or what should be done next (prescriptive analytics).
- Data analytics relies on a variety of software tools including spreadsheets, data visualization, reporting tools, data mining programs, and open-source languages for the greatest data manipulation.

UNDERSTANDING DATA ANALYTICS

Data analytics is a broad term that encompasses many diverse types of data analysis. Any type of information can be subjected to data analytics techniques to get insight that can be used to improve things. Data analytics techniques can reveal trends and metrics that would otherwise be lost in the mass of information. This information can then be used to optimize processes to increase the overall efficiency of a business or system.

For example, manufacturing companies often record the runtime, downtime, and work queue for various machines and then analyze the data to better plan workloads so the machines operate closer to peak capacity.





Data analytics can do much more than point out bottlenecks in production. Gaming companies use data analytics to set reward schedules for players that keep the majority of players active in the game. Content companies use many of the same data analytics to keep you clicking, watching, or re-organizing content to get another view or another click.

Data analytics is important because it helps businesses optimize their performances. Implementing it into the business model means companies can help reduce costs by identifying more efficient ways of doing business and by storing large amounts of data.

A company can also use data analytics to make better business decisions and help analyze customer trends and satisfaction, which can lead to new and better products and services.

DATA ANALYSIS STEPS

The process involved in data analysis involves several steps:

1. The first step is to determine the data requirements or how the data is grouped. Data may be separated by age, demographic, income, or gender. Data values may be numerical or divided by category.
2. The second step in data analytics is the process of collecting it. This can be done through a variety of sources such as computers, online sources, cameras, environmental sources, or through personnel.
3. The data must be organized after it's collected so it can be analyzed. This may take place on a spreadsheet or other form of software that can take statistical data.
4. The data is then cleaned up before analysis. It's scrubbed and checked to ensure that there's no duplication or error and that it is not incomplete. This step helps correct any errors before it goes on to a data analyst to be analyzed.

TYPES OF DATA ANALYTICS

Data analytics is broken down into four basic types:

- **Descriptive analytics:** This describes what has happened over a given period of time. Have the number of views gone up? Are sales stronger this month than last?
- **Diagnostic analytics:** This focuses more on why something happened. It involves more diverse data inputs and a bit of hypothesizing. Did the weather affect beer sales? Did that latest marketing campaign impact sales?
- **Predictive analytics:** This moves to what is likely going to happen in the near term. What happened to sales the last time we had a hot summer? How many weather models predict a hot summer this year?
- **Prescriptive analytics:** This suggests a course of action. We should add an evening shift to the brewery and rent an additional tank to increase output if the likelihood of a hot summer is measured as an average of these five weather models and the average is above 58%,

Data analytics underpins many quality control systems in the financial world, including the ever-popular Six Sigma program. It's nearly impossible to optimize something if you aren't properly measuring it, whether it's your weight or the number of defects per million in a production line.



The sectors that have adopted the use of data analytics include the travel and hospitality industry where turnarounds can be quick. This industry can collect customer data and figure out where problems, if any, lie and how to fix them.

Healthcare combines the use of high volumes of structured and unstructured data and uses data analytics to make quick decisions. Similarly, the retail industry uses copious amounts of data to meet the ever-changing demands of shoppers. The information that retailers collect and analyze can help them identify trends, recommend products, and increase profits.

The average total pay for a data analyst in the United States was just over \$80,000 in August 2023.

DATA ANALYTICS TECHNIQUES

Data analysts can use several analytical methods and techniques to process data and extract information. Some of the most popular methods include:

- **Regression analysis** entails analyzing the relationship between dependent variables to determine how a change in one may affect the change in another.
- **Factor analysis** entails taking a large data set and shrinking it into a smaller data set. The goal of this maneuver is to attempt to discover hidden trends that would otherwise have been more difficult to see.
- **Cohort analysis** is the process of breaking a data set into groups of similar data, often into a customer demographic. This allows data analysts and other users of data analytics to further dive into the numbers relating to a specific subset of data.
- **Monte Carlo simulations** model the probability of different outcomes happening. They're often used for risk mitigation and loss prevention. These simulations incorporate multiple values and variables and often have greater forecasting capabilities than other data analytics approaches.
- **Time series analysis** tracks data over time and solidifies the relationship between the value of a data point and the occurrence of the data point. This data analysis technique is usually used to spot cyclical trends or to project financial forecasts.

DATA ANALYTICS TOOLS

Data analytics has rapidly evolved in technological capabilities in addition to a broad range of mathematical and statistical approaches to crunching numbers. Data analysts have a broad range of software tools to help acquire data, store information, process data, and report findings.

Data analytics has always had loose ties to spreadsheets and Microsoft Excel. Data analysts also often interact with raw programming languages to transform and manipulate databases.

Data analysts also have help when reporting or communicating findings. Both Tableau and Power BI are data visualization and analysis tools used to compile information, perform data analytics, and distribute results via dashboards and reports.

Other tools are also emerging to assist data analysts. SAS is an analytics platform that can assist with data mining. Apache Spark is an open-source platform useful for processing large sets of data. Data analysts have a broad range of technological capabilities to further enhance the value they deliver to their company.

THE ROLE OF DATA ANALYTICS

Data analytics can enhance operations, efficiency, and performance in numerous industries by shining a spotlight on patterns. Implementing these techniques can give companies and businesses a competitive edge. The process includes four basic steps of analysis.

Data Mining

As the name suggests, this step involves “mining” or gathering data and information from across a broad spectrum of sources. Various forms of information are then recreated into the same format so they can eventually be analyzed. The process can take a good bit of time, more than any other step.

Data Management

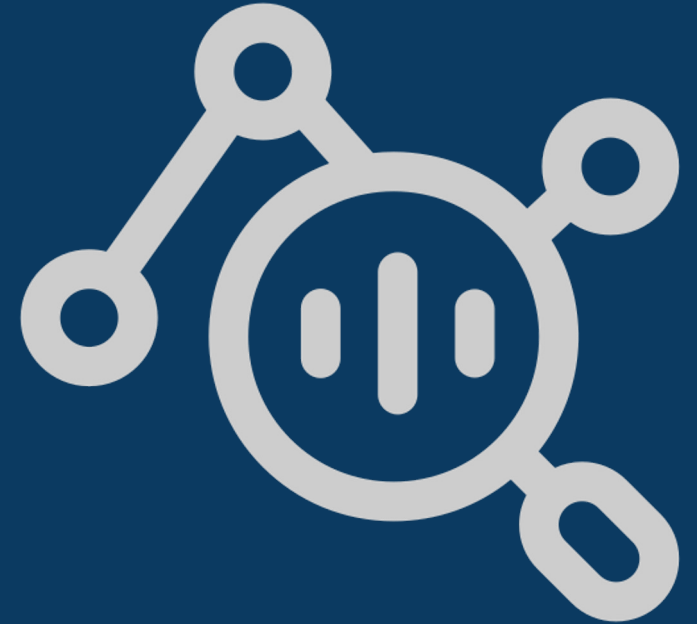
Data requires a database to contain, manage, and provide access to the information gathered through mining. The next step in data analytics is therefore the creation of such a database to manage the information. SQL was a common tool used for this purpose in the early days of data analytics and it's still widely used in 2023. Created in 1979, this computing language allows relational databases to be queried and the resulting data sets to be more easily analyzed.

Statistical Analysis

The third step is statistical analysis. It involves the interpretation of the gathered and stored data into models that will hopefully reveal trends that can be used to interpret future data. This is achieved through open-source programming languages such as Python. More specific tools for data analytics, like R, can be used for statistical analysis or graphical modeling.

Data Presentation

The results of the data analytics process are meant to be shared. The final step is formatting the data so it's accessible to and understandable by others, particularly those individuals within a company who are responsible for growth, analysis, efficiency, and operations. Having access can be beneficial to shareholders as well.



Importance and Uses of Data Analytics

Data analytics provide a critical component of a business's probability of success. Gathering, sorting, analyzing, and presenting information can significantly enhance and benefit society, particularly in fields such as healthcare and crime prevention. But the uses of data analytics can be equally beneficial for small enterprises and startups that are looking for an edge over the business next door, albeit on a smaller scale,

Why Is Data Analytics Important?

Implementing data analytics into the business model means companies can help reduce costs by identifying more efficient ways of doing business. A company can also use data analytics to make better business decisions.

What Are the 4 Types of Data Analytics?

Data analytics is broken down into four basic types. Descriptive analytics describes what has happened over a given period. Diagnostic analytics focuses more on why something happened. Predictive analytics moves to what is likely going to happen in the near term. Finally, prescriptive analytics suggests a course of action.

Who Is Using Data Analytics?

Data analytics has been adopted by several sectors where turnarounds can be quick, such as the travel and hospitality industry. Healthcare is another sector that combines the use of high volumes of structured and unstructured data, and data analytics can help in making quick decisions. The retail industry also uses large amounts of data to meet the ever-changing demands of shoppers.

The Bottom Line

Data analytics helps individuals and organizations make sure of their data in a world that's increasingly becoming reliant on information and gathering statistics. A set of raw numbers can be transformed using a variety of tools and techniques, resulting in informative, educational insights that drive decision-making and thoughtful management.



KAHOOT

MULTIPLE CHOICE QUESTIONS

[tps://create.kahoot.it/share/kahoot-para-la-evaluacion-formativa/e32fa2af-7393-4463-8b73-eea564f5e2e2](https://create.kahoot.it/share/kahoot-para-la-evaluacion-formativa/e32fa2af-7393-4463-8b73-eea564f5e2e2)



USEFUL INTERNET VOCABULARY



TIC

BROWSER

A computer program with a graphical user interface for displaying and navigating between web pages.

SERVER

A person or thing that serves.

SPAM

Any kind of unwanted, unsolicited digital communication that gets sent out in bulk

LIVE STREAM

A live transmission of an event over the internet.

CSS

CSS is a language for specifying how documents are presented to users

WI-FI

A wireless networking technology that uses radio waves to provide wireless high-speed Internet access

DISCUSSION QUESTIONS ABOUT NLP

--- FORUM PARTICIPATION ---

- What is Natural Language Processing (NLP), and how does it relate to artificial intelligence?
- Discuss real-world applications of NLP. How is it being used in various industries, and what impact does it have on daily life?
- Explore the challenges associated with NLP, such as ambiguity and bias. How do these challenges affect the development and deployment of NLP systems?

USEFUL NATURAL LANGUAGE PROCESSING (NLP) TERMINOLOGY

TOKENIZATION

is the process of breaking down a text into smaller units called tokens. These tokens can be words, phrases, or sentences, depending on the level of granularity needed.

STEMMING

is a technique used to reduce words to their base or root form. It involves removing suffixes or prefixes from words, so variations of a word are represented by the same stem.

PART-OF-SPEECH TAGGING

involves assigning a specific part-of-speech (such as noun, verb, adjective, etc.) to each word in a sentence.

WORD EMBEDDINGS

are dense vector representations of words in a continuous vector space. Each word is assigned a unique vector, and the distance and direction between these vectors capture semantic relationships between words.

AMBIGUITY IN NLP

the presence of multiple possible interpretations or meanings in a given language context.

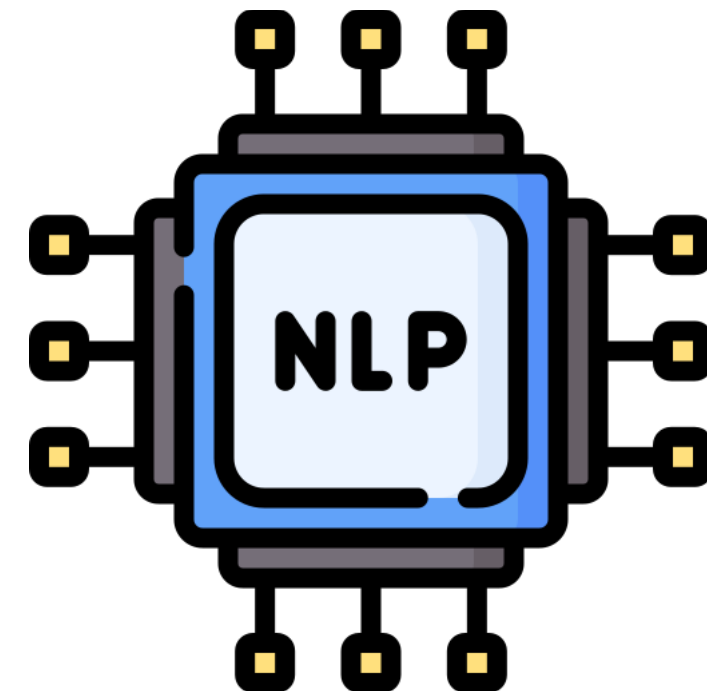
BIAS IN NLP

systematic and unfair preferences or prejudices in language processing models.

Natural Language Processing (NLP): Revoluzionating Human-machine communication.

Natural Language Processing (NLP) is a dynamic and rapidly advancing field within computer science that focuses on the interaction between computers and human languages. It encompasses a wide range of techniques and applications that enable computers to understand, analyze, and generate human language, playing a pivotal role in facilitating seamless communication and interaction between humans and machines.

NLP has made significant strides in various domains, including machine translation, speech recognition, text analysis, and question answering, and continues to drive innovation across industries. It involves the study and development of computational models and algorithms that empower computers to process and comprehend natural language, bridging the gap between human communication and machine understanding.



NLP comprises several subfields, each with its own set of challenges and applications:

Speech Recognition: this subfield focuses on developing systems that enable computers to convert spoken language into written text or commands. Speech recognition systems have found practical applications in voice-activated assistants, dictation software, call centers, and various voice-controlled devices, enhancing accessibility and efficiency in human-computer interactions.

Machine Translation: machine translation aims to automatically translate text or speech from one language to another. With the rise of globalization and the need for multilingual communication, machine translation systems have become increasingly crucial. They facilitate cross-language communication in areas such as international business, travel, content localization, and cultural exchange.

Text Analysis: text analysis focuses on extracting meaningful information and insights from written text. It encompasses various tasks, including sentiment analysis, topic modeling, named entity recognition, and document classification. Text analysis techniques are widely employed in social media monitoring, customer feedback analysis, market research, content recommendation systems, and information retrieval.

Question Answering:

question-answering systems aim to comprehend and respond to questions posed in natural language. These systems employ techniques such as information retrieval, natural language understanding, and knowledge representation to provide accurate and relevant answers. Question-answering systems have applications in customer support, educational platforms, and research domains where quick and precise information retrieval is essential.



The history of NLP traces back to the early days of artificial intelligence (AI). In the 1950s, researchers began exploring the idea of using computers to process and understand human language. One notable milestone was the development of the ELIZA program in the 1960s, which simulated a conversation with a human therapist using simple pattern-matching techniques. ELIZA laid the foundation for subsequent advancements in NLP.

Throughout the 1970s and 1980s, NLP research advanced significantly, incorporating statistical methods and machine learning techniques. These developments improved the accuracy and flexibility of NLP systems, enabling real-world applications in machine translation, speech recognition, and information retrieval.

The emergence of the World Wide Web in the 1990s fueled a renewed interest in NLP, as vast amounts of textual data became readily accessible. Researchers explored techniques such as web scraping, text mining, and document clustering to extract valuable insights from the ever-expanding online content, opening up new possibilities for NLP applications.



The 2000s witnessed a revolutionary phase in NLP with the advent of deep learning techniques. Deep learning, a subset of machine learning, utilizes neural networks with multiple layers to learn hierarchical representations of data. This approach significantly improved the performance of NLP systems, leading to breakthroughs in machine translation, text classification, sentiment analysis, and natural language generation. The application of deep learning in NLP has also facilitated the development of large-scale language models capable of generating coherent and contextually relevant human-like text.

NLP finds extensive applications across various industries and domains:

Healthcare: NLP assists in extracting crucial information from medical records, clinical notes, and research articles, aiding in diagnosis, treatment recommendation, and drug discovery. It enhances healthcare professionals' ability to access and analyze vast amounts of patient data efficiently.

Customer Service: NLP-powered chatbots and virtual assistants provide automated support and address customer queries in a conversational manner. These systems offer personalized assistance, improving customer satisfaction, reducing response times, and freeing up human agents for more complex tasks.

Social Media Analysis: NLP techniques are used to monitor and analyze social media data, enabling sentiment analysis, trend detection, and brand reputation management. Businesses and organizations leverage these insights to understand customer preferences, identify emerging trends, and tailor their marketing strategies accordingly.

E-commerce: NLP systems improve product search and recommendation systems, enabling personalized shopping experiences. They analyze customer reviews and feedback, facilitating sentiment analysis and enhancing product quality and customer satisfaction.

Financial Services: NLP is employed in tasks such as fraud detection, risk assessment, and automated document analysis. By extracting critical information from financial documents, NLP systems enable efficient decision-making, regulatory compliance, and improved customer service in the financial sector.

Looking ahead, the future of NLP holds tremendous potential for further advancement and innovation. Exciting developments to watch for include:

Contextual Understanding: NLP systems are evolving to better understand the context in which language is used. This includes comprehending idioms, sarcasm, and nuanced meanings, leading to more accurate and human-like responses. Such advancements will contribute to more natural and engaging human-machine interactions.

Multimodal NLP: the integration of NLP with other modalities, such as images, videos, and sensor data, enables a more comprehensive understanding of the world. Multimodal NLP systems will provide richer and more interactive user experiences, with applications ranging from virtual reality to autonomous vehicles.

Ethical and Responsible NLP: as NLP becomes more pervasive, addressing ethical concerns such as bias, fairness, and privacy will be crucial. Researchers and practitioners are actively working on developing frameworks and techniques to ensure the responsible and unbiased use of NLP technologies, promoting transparency and accountability.

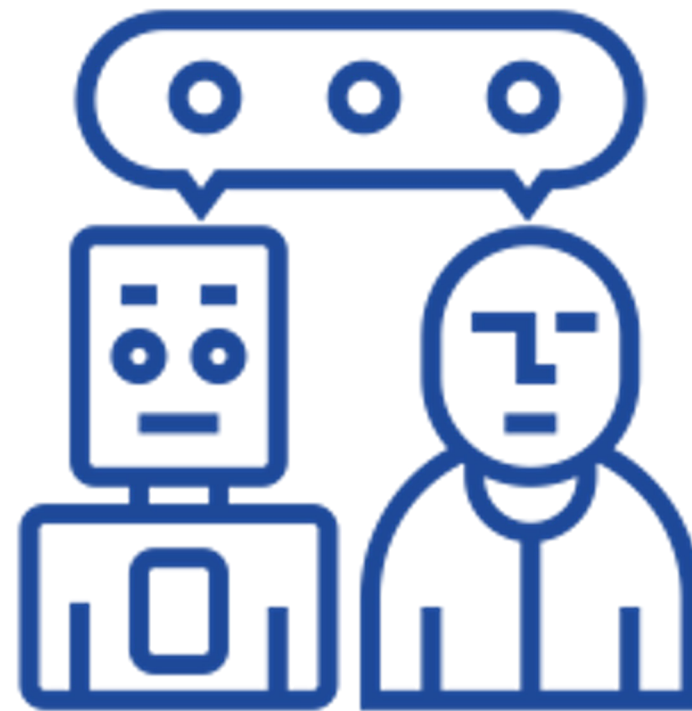
Low-Resource Languages: Efforts are being made to improve NLP capabilities for languages with limited resources and linguistic datasets. This includes developing methods to transfer knowledge from high-resource languages to low-resource ones, enabling wider accessibility and inclusivity.

Domain-Specific NLP: Tailoring NLP systems to specific domains, such as legal, medical, or scientific, will lead to more accurate and specialized applications. Domain-specific NLP models can better understand domain-specific terminology and context, enhancing performance and reliability in niche areas.

In conclusion, NLP has witnessed remarkable progress and has had a significant impact on various industries and everyday life. With ongoing research, advancements in deep learning, and the exploration of new frontiers such as multimodal understanding and ethical considerations, the future of NLP looks incredibly promising. As NLP continues to evolve, we can expect a world where human and machine interactions through natural language become increasingly seamless, effective, and aligned with our communication needs.

Taken from:

<https://medium.com/@mubarakdaha/introduction-to-natural-language-processing-nlp-revolutionizing-human-machine-communication-ba06d186210d>



MULTIPLE CHOICE

QUIZZZ



WHAT IS NATURAL LANGUAGE PROCESSING (NLP)



<https://www.youtube.com/watch?v=CMrHM8a3hqw>