

# Carlos Andrés Betancur Yepes.

**Physicist: Data Scientist | Data Specialist.**

Phone: +57 3207712588 | E-mail: [ndrsbtncr@gmail.com](mailto:ndrsbtncr@gmail.com)

Trained scientist with strong skills in **statistical** and **computational** interpretation of industry problems, backed by my undergraduate degree in **Physics**, where I gained a foundation in analytical methods and programming, especially in **Python**. I have also completed courses in **SQL**, **Power BI**, **Power Apps** and **Power Automate** thus strengthening my skills in **data analysis**. As a **data scientist**, I have expanded my programming knowledge with a focus on libraries such as **Pandas**, **JSON**, **Numpy**, **Matplotlib**, **Seaborn**, **NLTK**, and in the field of **Machine Learning** where I have applied **Azure services**, **supervised** and **unsupervised** models from **Scikit-Learn**, and **natural language processing** models using **Transformers**, including the integration of the **Chat GPT API**. I also have experience in **Deep Learning** tools like **Tensorflow** and **Pytorch**. Furthermore, I have cultivated skills in **web development** using **Flask**, **Streamlit**, **HTML**, **CSS**, and **JavaScript** to deploy models in production. I stand out for my ability to learn quickly, **collaborate efficiently in teams**, and ensure effective deliveries. I am committed to excellence and continuous improvement, seeking opportunities to apply my **analytical skills** and contribute to the **success of the team and organization**.

## EXPERIENCE

### Data Scientist at Subocol (April 2024 – Present)

- I specialize in automating time-consuming manual tasks, optimizing processes, and improving the company's operational efficiency. I use **Python** to design and develop **automated solutions** that significantly reduce the manual workload, allowing the team to focus on higher-value strategic activities. My goal is to increase productivity and accuracy in task execution, contributing to the overall success of the organization.

### Data Scientist at ManpowerGroup Colombia (December 2023 – March 2024)

- I implemented and optimized an **Azure AI Search service** for intelligent document management. I developed an innovative user interface using **Python**, **Streamlit**, **HTML**, and **CSS**, significantly enhancing efficiency in information retrieval. This project underscored my ability to tackle complex challenges and deliver effective solutions in **artificial intelligence** and data

management, positioning the organization at the forefront of technological innovation.

- I developed a customer experience comment **classification model**, where I implemented a **bagging-type** model using **machine learning** algorithms. Additionally, the **ChatGPT API** was utilized to enhance accuracy in categorizing comments into various categories. This project achieved a significant improvement in the efficiency of the classification process, receiving positive feedback for enhancing the understanding of comments.
- Use of **Python** to consume **APIs** in order to extract information through **Web Scraping** techniques, utilizing the **requests** and **Beautiful Soup** libraries. Implementation of an automated process for data extraction, with manipulation of documents in **JSON** format and unstructured text. The results are automatically incorporated into a Word file with a carefully formatted layout for easy reading by the user.

### **Data Analyst at ManpowerGroup Colombia (January 2023 – December 2023)**

- I restructured a crucial **Power BI dashboard** for the company, focusing on operational costs and reducing the update time **from over 8 hours to just 3 minutes**. Through efficient data modeling and DAX programming, along with optimization of the **ETL in Power Query**, I streamlined access to key financial information, significantly enhancing the company's ability to make informed decisions regarding costs and operations.
- Development and Automation of **Power BI** dashboards integrated with tools such as **Python** and **SQL**. I also consumed **APIs** through **Power BI** using queries in M language.
- Optimizing queries by applying advanced **Power Query** and **M language** usage for big data projects.
- Building advanced formulas in **DAX language** and designing interactive reports leveraging the various functionalities in **Power BI** visuals for **descriptive** and **predictive analytics**.

## **MY INDEPENDENT PROJECTS**

**About data mining, machine learning, and web scraping:**

[https://github.com/ndrsbtncr/My\\_Projects\\_ML](https://github.com/ndrsbtncr/My_Projects_ML)

- In my repository, you can find projects about **natural language processing** such as **sentiment analysis** (NLTK, SpaCy, SciPy, TextBlob), **machine learning models** (Linear Regression, Decision Trees, Random Forest, Logistic Regression, Kmeans, Kprototypes), and **anomaly detection** related

to cybersecurity (IsolationForest). I apply **probabilistic techniques** for device **survival analysis** and predict failure events. I've also created a Python and Selenium **Web Scraping** robot with an **executable .exe** for data extraction without using **APIs**.

## **MATERIALS SCIENCE RESEARCH GROUP, Universidad de Antioquia (June 2021 – December 2022)**

- **Development of new theoretical studies** on optical layers of experimentally verified non-linear materials.
- **Implementation of news algorithms** for the solution of differential equations in Wolfram Mathematica language.

## **SKILLS.**

- Python, Machine Learning, Deep Learning, C++, HTML, CSS, JavaScript, SQL, Power BI, Microsoft Excel, Mathematical and statistical modeling, Power Pivot, Power Query, M language, Power Apps, Power Automate, Data Mining, Postman, Matlab, Wolfram Mathematica, intermediate English, basic German, Adaptability, Effective Communication, Continuous Learning, Innovation, Teamwork, Leadership.

## **EDUCATION**

### **PROFESSION**

#### **Universidad de Antioquia**

- **Physics** (2014-2022)

**Achievement:** I received an honor mention in the ninth semester as the physics student with the highest-grade average.

### **CERTIFICATIONS**

#### **Universidad de Antioquia**

- Big Data (April 2022 - November 2022)

#### **MinTic**

- Fundamentals of data analytics (November 2022 – January 2023)

## Udemy

- Basic, Intermediate, and Advanced Excel. (December 2022 - December2022)
- SQL Server from beginner to advanced. (January 2023 – February 2023)
- Power BI from beginner to advanced. (January 2023 – April 2023)

## REFERENCES

Senior Data Scientist Michal Hemmerling

- E-mail: [Michal.hemmerling@udea.edu.co](mailto:Michal.hemmerling@udea.edu.co)
- Phone1: +55 16988511933
- Phone2: +55 16988067251

Physicist Lucas Quiceno

- E-mail: [lucas.quiceno@udea.edu.co](mailto:lucas.quiceno@udea.edu.co)
- Phone: 3024611660

## CONTACT

Phone: +57 3207712588

E-mail: [ndrsbtncr@gmail.com](mailto:ndrsbtncr@gmail.com)

Linkedin: <https://www.linkedin.com/in/candres-betancur-data-scientist/>

Birthplace: Frontino-Antioquia

Birthdate: 26/11/1995

Address: Copacabana, Antioquia