

Practical Session

Data Mining & Neural Networks

Dr. Wilmer Garzón

Director, Master's Program in Data Science
Department of Computer Engineering



Escuela Colombiana de Ingeniería
Universidade da Coruña

2025



UNIVERSIDAD

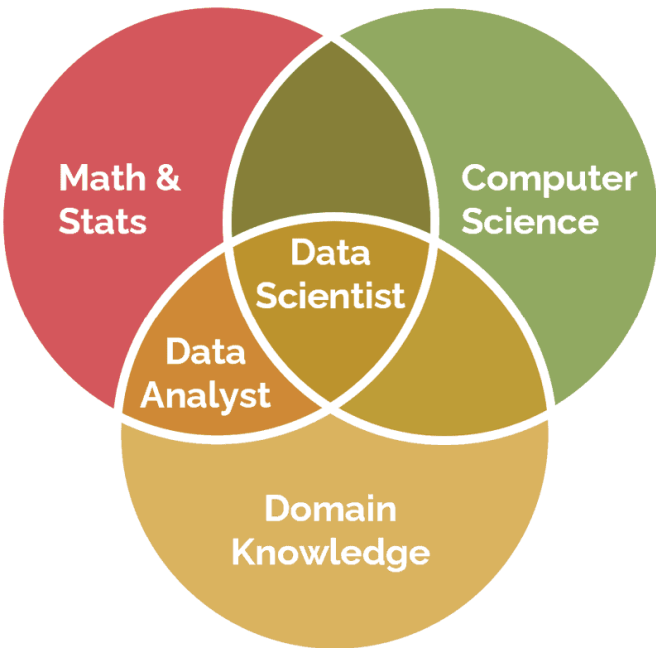
Tools We Will Use

- **Python:** Our primary programming language
- **Jupyter Notebooks:** Interactive coding environment
 - Local Jupyter (via Anaconda)
 - Google Colab (cloud-based, no installation)

```
listsum.py
1 # The sum of Integers,
2 # sum of square and cube of integers
3 """The sum of 'n' Integers, sum of squares, cubes"""
4 # Input for the end number, 'n'
5 nint = int(input('Enter the end Integer value: '))
6 #
7 # sum of integers up to integer 'n'
8 nsum=0
9 nsqsum = 0
10 ncusum = 0
11 for i in range(nint):
12     count=i+1
13     nsum = nsum + count
14     nsqsum = nsqsum + count* count
15     ncusum = ncusum + count * count * count
16 print('Sum of n integers = ', nsum)
17 print('Sum of square of n integers = ', nsqsum)
18 print('Sum of cube of n integers = ', ncusum)
19 print('ncube sum is the same as square of nsum ', nsum*nsum)
20
```

PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
PS C:\Users\itish\VScode-Dir> python listsum.py
Enter the end Integer value: 100
Sum of n integers = 5050
Sum of square of n integers = 338350
Sum of cube of n integers = 25502500
ncube sum is the same as square of nsum 25502500
C:\Users\itish\VScode-Dir>
```



Why Python?

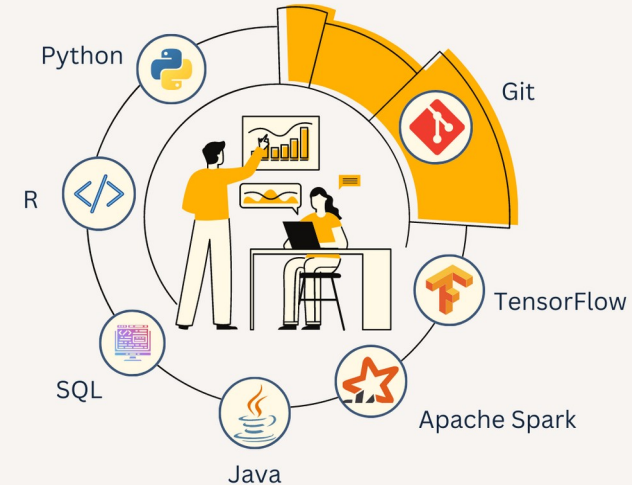


- Python's Prominence:
 - Widely embraced in the realms of data science and machine learning.
 - Renowned for its simplicity and ease of use.
 - Boasts a comprehensive ecosystem of libraries and tools that enhance functionality.

Key Libraries

- **NumPy:** Numerical computing
- **Pandas:** Data manipulation
- **Matplotlib / Seaborn:** Data visualization
- **Scikit-learn:** Machine learning algorithms
- **TensorFlow / Keras:** Deeplearning frameworks

Top 7 data science tools you must learn



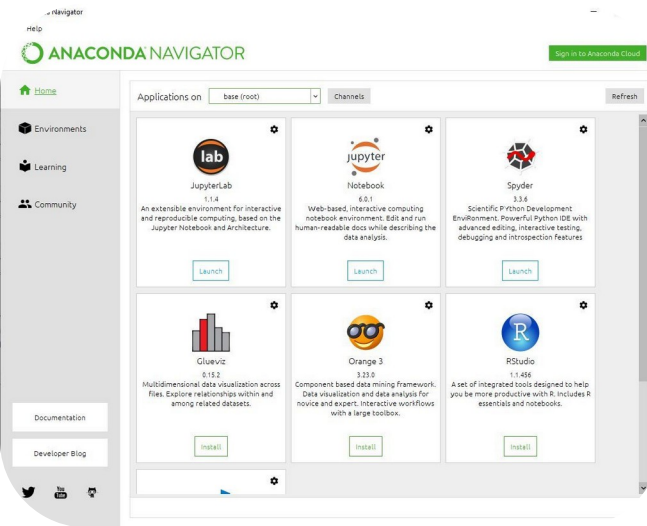
Practical Sessions

- Daily hands-on labs
- Real-world datasets
- Guided exercises and challenges
- Optional collaborative work



Hands-On Labs

A Virtual Hands-on Lab Experience



Getting Started

- Install Anaconda for local Jupyter
- Set up a Google account for Colab
- Download initial notebook from the course

Questions?

Ask about tools, setup, or course expectations
We're here to support your learning journey!

Introduction to Python

- **High-level, interpreted language** that is designed for ease of use and flexibility
- **Known for readability and simplicity**, making it accessible for beginners and experts alike
- **Widely used in data science and AI**, facilitating complex computations and data analysis tasks



```
listsum.py
1  # The sum of Integers,
2  # sum of square and cube of integers
3  """The sum of 'n' Integers, sum of squares, cubes"""
4  # Input for the end number, 'n'
5  nint = int(input('Enter the end Integer value: '))
6  #
7  # sum of integers up to integer 'n'
8  nsum=0
9  nsqsum = 0
10 ncusum = 0
11 for i in range(nint):
12     count=i+1
13     nsum = nsum + count
14     nsqsum = nsqsum + count* count
15     ncusum = ncusum + count * count * count
16 print('Sum of n integers = ', nsum)
17 print('Sum of square of n integers = ', nsqsum)
18 print('Sum of cube of n integers = ', ncusum)
19 print('ncube sum is the same as square of nsum ', nsum*nsum)
20
```

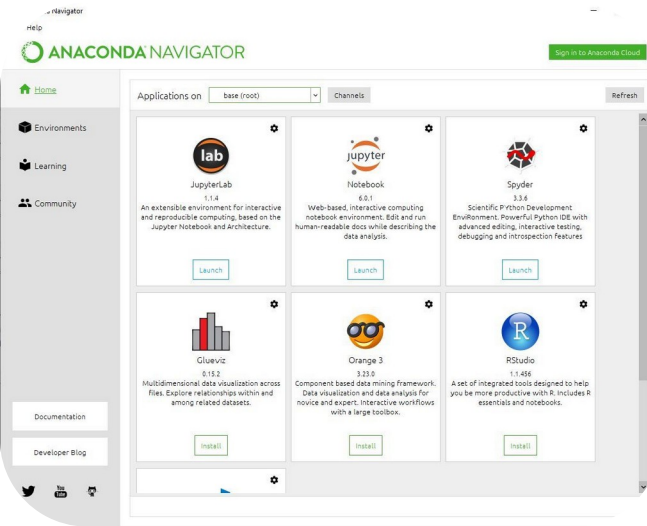
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS

```
PS C:\Users\itish\VScode-Dir> python listsum.py
Enter the end Integer value: 100
Sum of n integers = 5050
Sum of square of n integers = 338350
Sum of cube of n integers = 25502500
ncube sum is the same as square of nsum 25502500
PS C:\Users\itish\VScode-Dir>
```



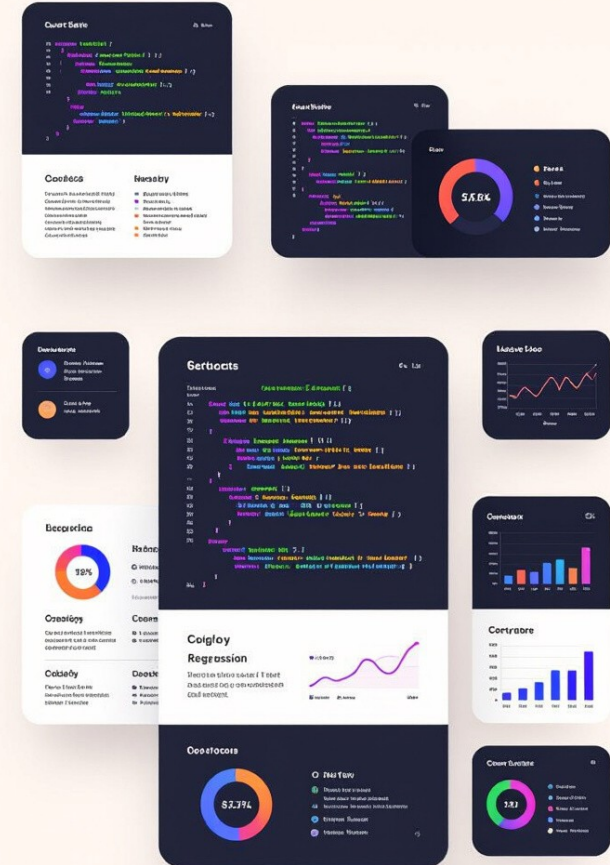
Setting Up Python Environment

- **Install Anaconda** (includes Python and Jupyter)
- **Create virtual environments**
- **Use pip or conda** to install libraries



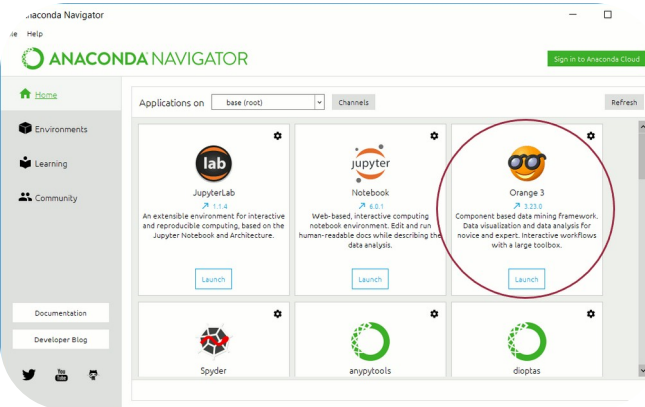
Jupyter Notebooks (JpNb)

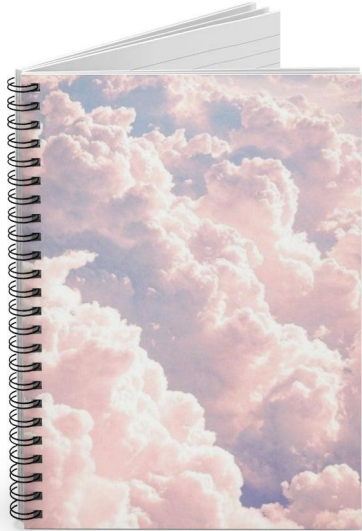
- **Combine code, text, and visualizations** to create comprehensive and informative documents that enhance understanding.
- **Ideal for data exploration and documentation** as it allows users to present findings clearly and effectively.
- **Supports interactive learning** by enabling users to engage with the material in a dynamic and meaningful way.



Using Local JpNb

- Launch the application conveniently via **Anaconda Navigator** or through the terminal interface.
- Save your notebooks locally on your device for easy access and management.
- The execution resources are associated with your computing machine.
- Enjoy full control over your environment and the various packages you choose to install and utilize.





Using Google Colab

- **Access notebooks** from any device, ensuring flexibility and convenience for all users
- **No installation required**, allowing for immediate use without any setup hassles
- **Free access to GPUs** for faster training, significantly enhancing computational efficiency and performance
- The data is not stored in the cloud, only is temporal

Array Operations

- Operations may combine two arrays if they have exactly the same length.

```
x=[1 3 5 7]
```

```
y=[10 20 30 40]
```

```
x+y
```

```
y-x
```

- The operations $+$ and $-$ work as expected.
- For element-by-element multiplication, division, and exponentiation, use `.*`, `./`, and `.^`.

(Explanation: The usual operators `*`, `/`, and `^` perform matrix [Linear algebra] operations between arrays. We will not cover that in CS100.)

Introduction to NumPy



- **Efficient array operations** that optimize performance and resource usage
- **Supports linear algebra** and statistics, facilitating complex calculations and data analysis
- **Foundation for many other libraries** that build upon its capabilities, enhancing functionality and versatility.

Introduction to Pandas



- **Work with tabular data** using DataFrames to efficiently manage and analyze structured information
- **Easy data cleaning** and transformation processes streamline the preparation of datasets for analysis
- **Powerful grouping** and aggregation tools enable comprehensive insights and summaries of complex data sets.

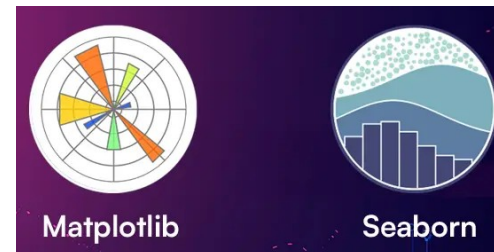
Columns

rows

Regd. No	Name	Marks%
1000	Steve	86.29
1001	Mathew	91.63
1002	Jose	72.90
1003	Patty	69.23
1004	Vin	88.30

Introduction to Matplotlib and Seaborn

- Create detailed plots and informative charts that effectively communicate data insights
- **Seaborn** significantly enhances **Matplotlib** by providing better visuals and aesthetics for data representation
- This tool is particularly useful for thorough data exploration and compelling storytelling through visual means.





Introduction to Scikit-learn

- A simple and intuitive interface for various machine learning models, making it user-friendly for all levels of expertise
- Includes essential techniques such as classification, regression, and clustering to cover a wide range of applications
- Comprehensive tools for model evaluation and preprocessing, ensuring high-quality results and efficient workflows.

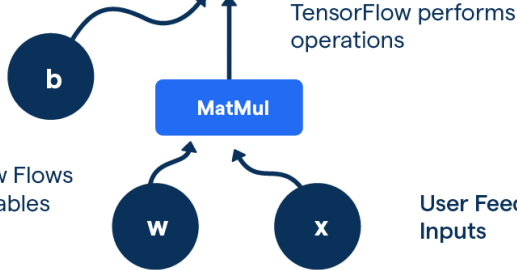


Tensorflow



User Fetches Outputs

TensorFlow Flows
tensors



Introduction to TensorFlow and Keras

- **TensorFlow:** A powerful and versatile deep learning framework that facilitates complex computations and model training.
- **Keras:** An intuitive and user-friendly API designed specifically for building and training neural networks efficiently.
- This technology is widely used for various applications, including image recognition, text processing, and sequence modeling tasks.



Keras



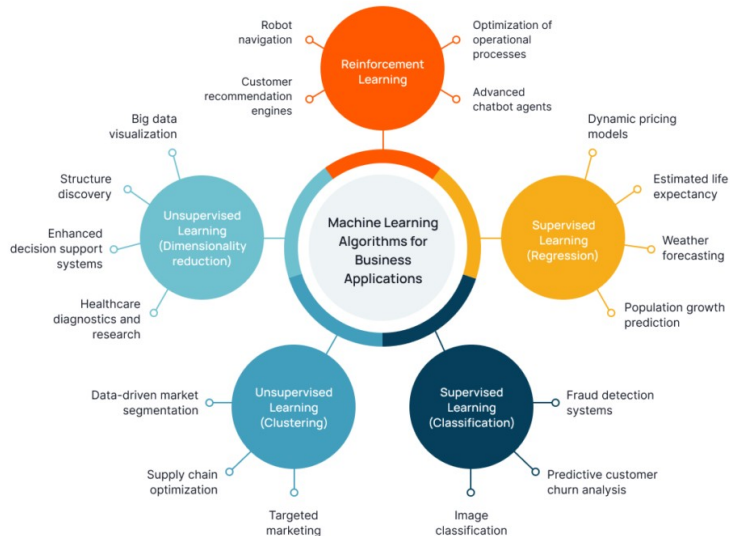
TensorFlow

2.0

Real-world Applications

intelliarts

Machine Learning Algorithms for Business Applications



Healthcare: disease prediction technologies help identify potential health risks early on

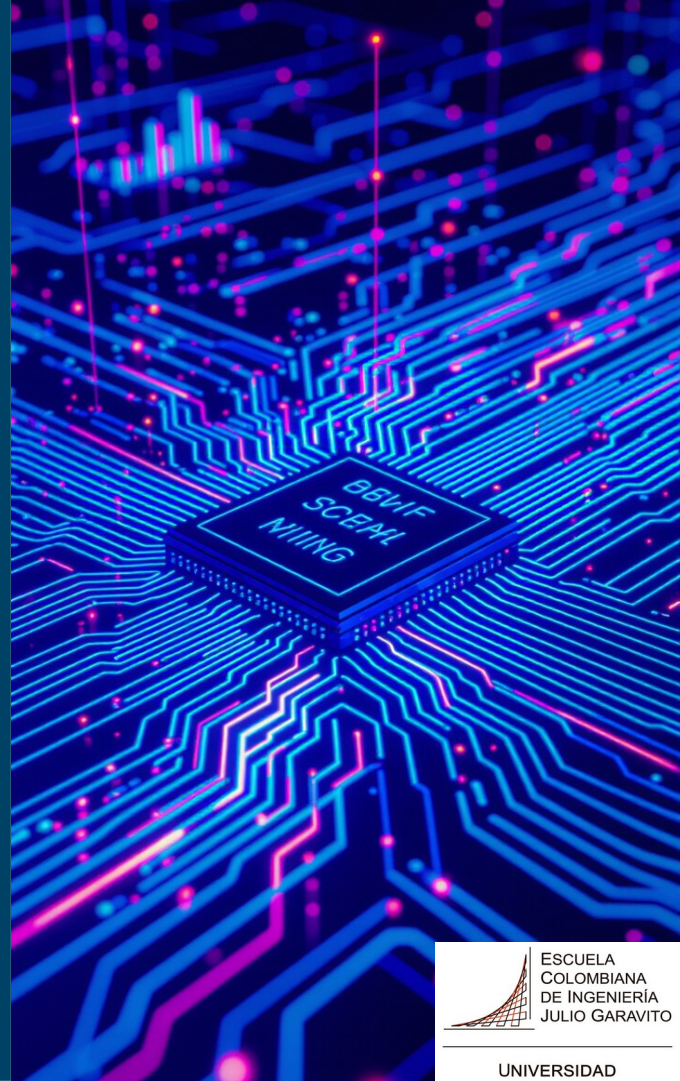
Finance: fraud detection systems are crucial for safeguarding financial transactions

Marketing: customer segmentation allows businesses to tailor their strategies effectively

Projects will reflect real-world use cases, demonstrating practical applications and benefits in various industries.

Case Guidelines

- Choose a topic related to **data mining** or **neural networks**
- Use **real datasets**
- Follow provided structure and milestones
- Present findings at the end of the case



Tips for Success

- **Practice and Practice** regularly
- Don't hesitate to ask for help
- Collaborate and share ideas
- Explore beyond the course materials

