**“BioInformatics”**

**Course**

**a.a. 2020-2021**

**Coronavirus’ Study #7**

|  |  |  |
| --- | --- | --- |
| **Team Members** | | |
| **Name& Surname** | **Matriculation Number** | **E-mail address** |
| **Alessandro Pio** | *261764* | alessandro.pio@student.univaq.it |
| **Lorenzo Tucceri Cimini** | *259039* | lorenzo.tuccericimini@student.univaq.it |
| **Lorenzo Negro** | *261144* | lorenzo.negro@student.univaq.it |

A. Pre-installation

It is mandatory to install and configure some tools in order to run our projects correctly.  
The steps to be performed are:

1. Download PyCharm (<https://www.jetbrains.com/pycharm/download/>)
2. Configure Pycharm
3. Download ClustalO (<http://www.clustal.org/omega/#download>)
4. Installing PyCharm

PyCharm is an integrated development environment(IDE) used in computer programming, specifically for the Python language.

Python is a general-purpose language, used by data scientists and developers, which makes it easy to collaborate across organizations through its simple syntax.

* Go to PyCharm download page:

Immagine che contiene testo

Descrizione generata automaticamente

* Select the installation file according toyour operating system (Windows, macOS, Linux).
* Download the file.
* Run it.

1. ConfigurePyCharm

* Open the IDE.
* Import our project by selecting*File > Open* and then browse to the folder containing it.

Immagine che contiene testo, segnale, dispositivo, screenshot

Descrizione generata automaticamente

* Configure our project’s interpreter by selecting *File > Settings* and then *Python 3.8* (or higher version).  
  IDE helps downloading the wanted Python version if it is not already available.

Immagine che contiene testo

Descrizione generata automaticamente

1. ClustalO

# ClustalO is an essential tool in our project: it allows performing a Multiple Sequence Alignment.

* Download ClustalO from the link below:

(<http://www.clustal.org/omega/#download>)

Immagine che contiene testo

Descrizione generata automaticamente

* Select the binary according to your OP.

B. Running Program

Now that the pre-installation is completed, it is time to run our project by using the arrow shown in the ribbon located on the top-left

Immagine che contiene testo, orologio

Descrizione generata automaticamente

or using the shortcut (CapsLock + F10).

The “python console” allows the user to insert needed data.

Immagine che contiene testo

Descrizione generata automaticamente

**PROJECT 1**

This program implements a pipeline,which identifies sequences located in a “.fasta” format file and produces an output of:

* ID of thesequence;
* Lengthof thesequence;
* Number of mutations in the sequence.

Immagine che contiene testo

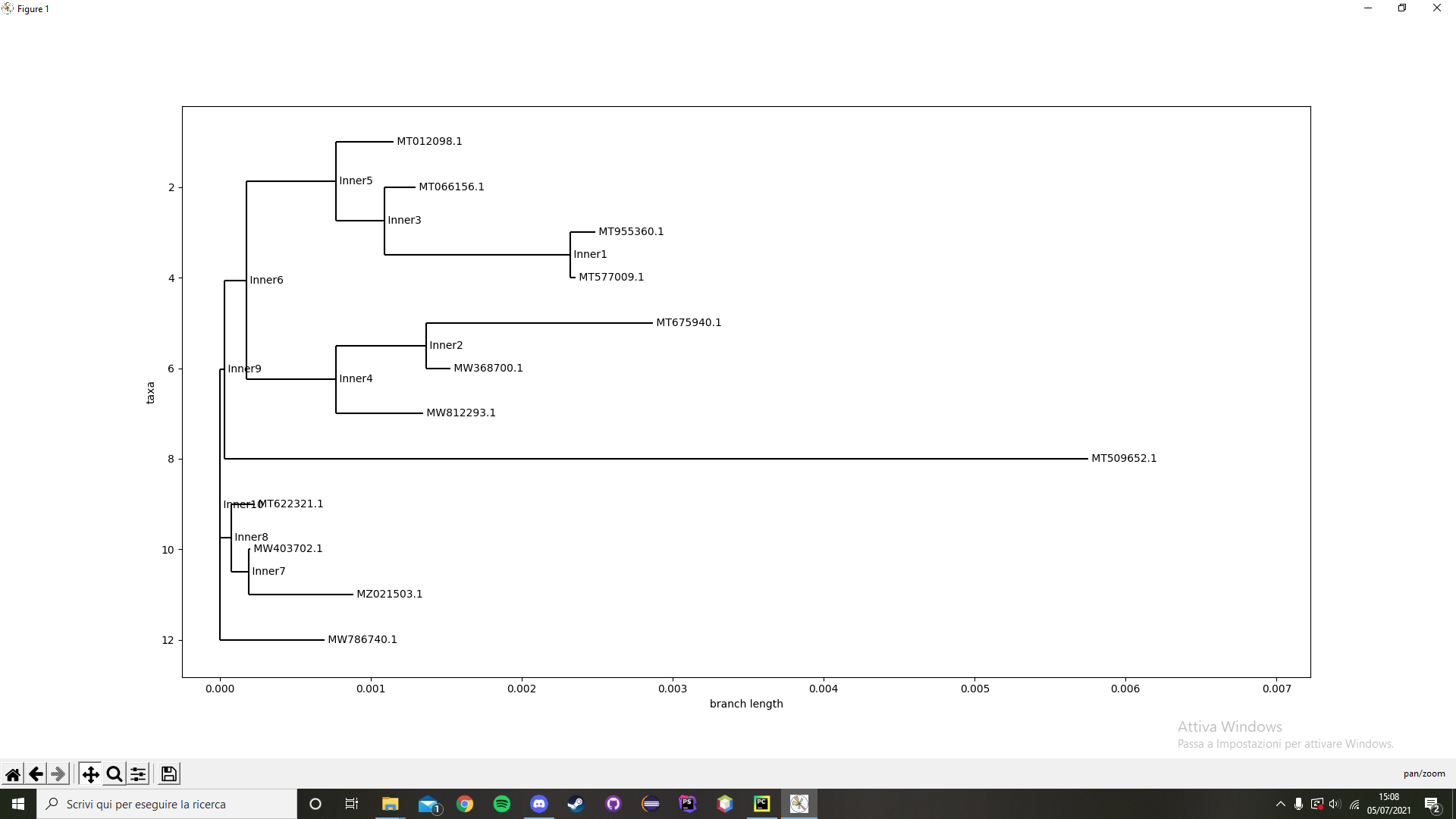
Descrizione generata automaticamente

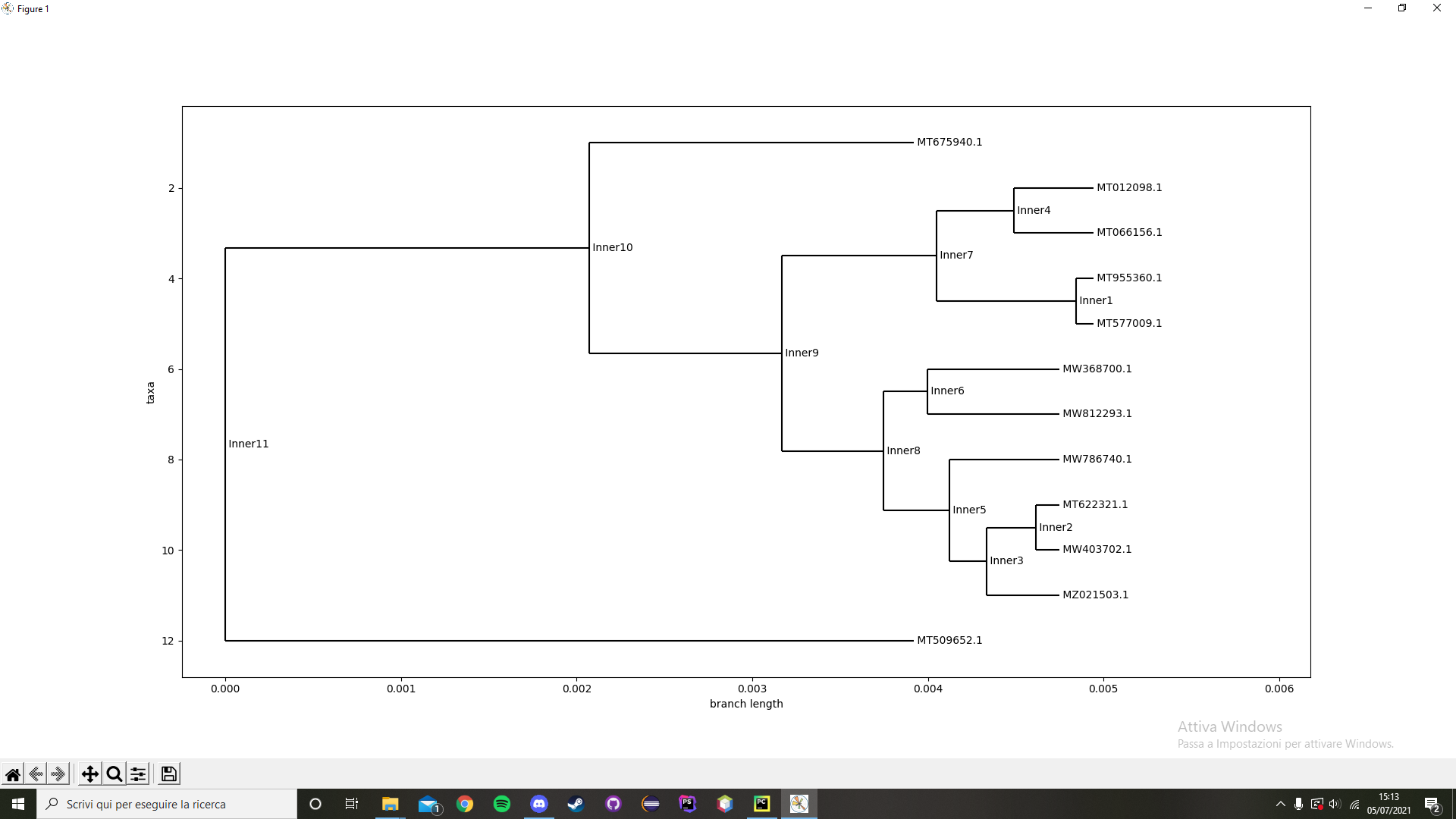
**PROJECT 2**

This program implements a pipeline, which builds phylogenetic trees of 2 clusters of Covid19 sequences.

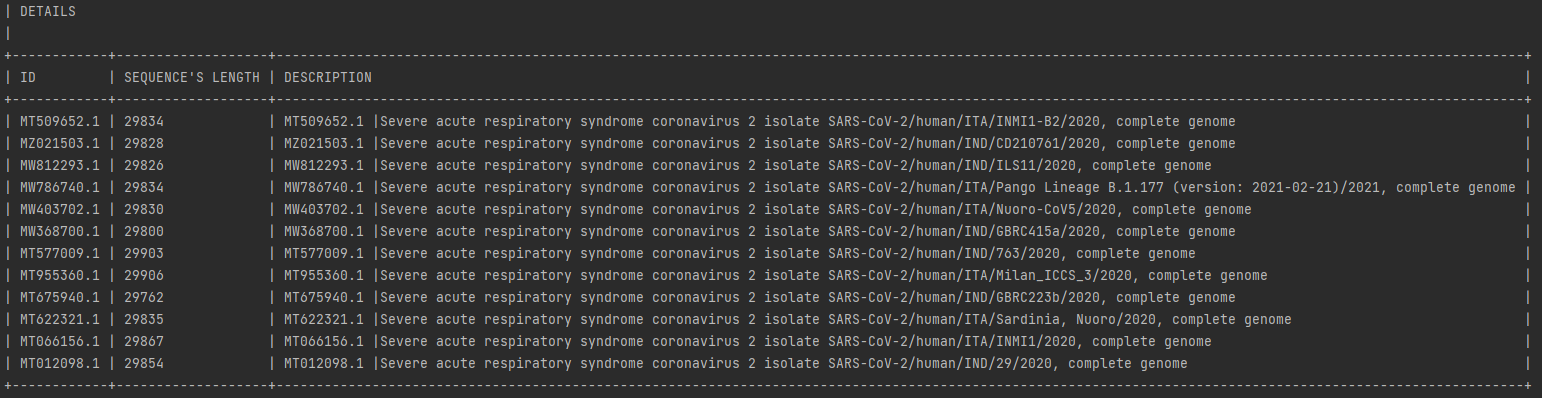
Two types of trees are built:

* Neighbor Join Tree
* UPGMA Tree

(Neighbor Join Tree)



(UPGMA Tree)



(Legend for Tree’s)