**“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. Installazioni generali

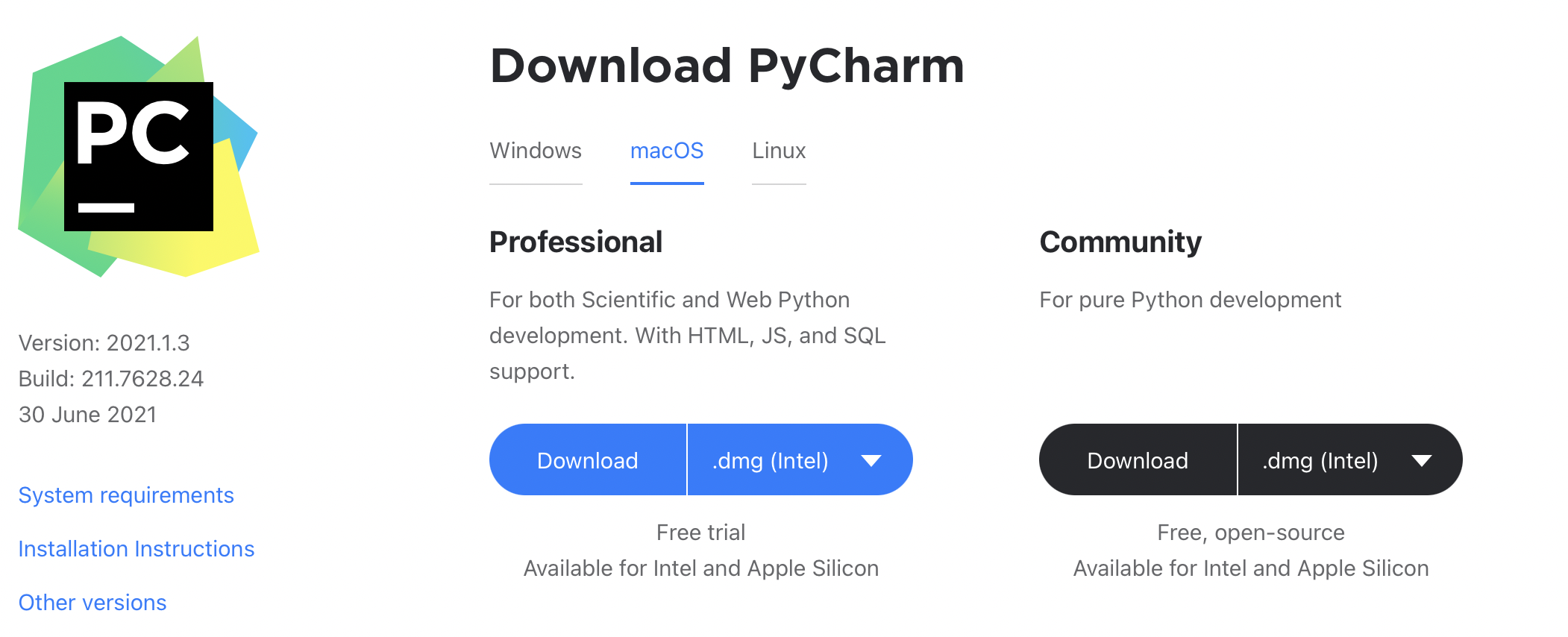
Prima di procedere all’esecuzione del software, bisogna installare i seguenti tool che sono indispensabili per il suo corretto funzionamento:

1. PyCharm (<https://www.jetbrains.com/pycharm/download/>)
2. Configure Pycharm
3. Create Configurations
4. Clustal Omega (<http://www.clustal.org/omega/#download>)
5. (Opzionale) Installazione dei package importanti per l’avvio dei software
6. Installing PyCharm

Pycharm è un ambiente di sviluppo integrato (IDE) utilizzato nella programmazione dei software, in particolare con il linguaggio Python.

Python è un linguaggio generico, utilizzato dagli scienziati dei dati e dagli sviluppatori, il che rende facile collaborare attraverso la sua semplice sintassi

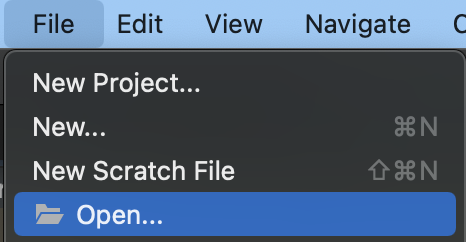
Per installare questo software bisogna recarsi nella pagina di download



A seconda della architettura utilizzata, si potrà scegliere il file di installazione adeguato.

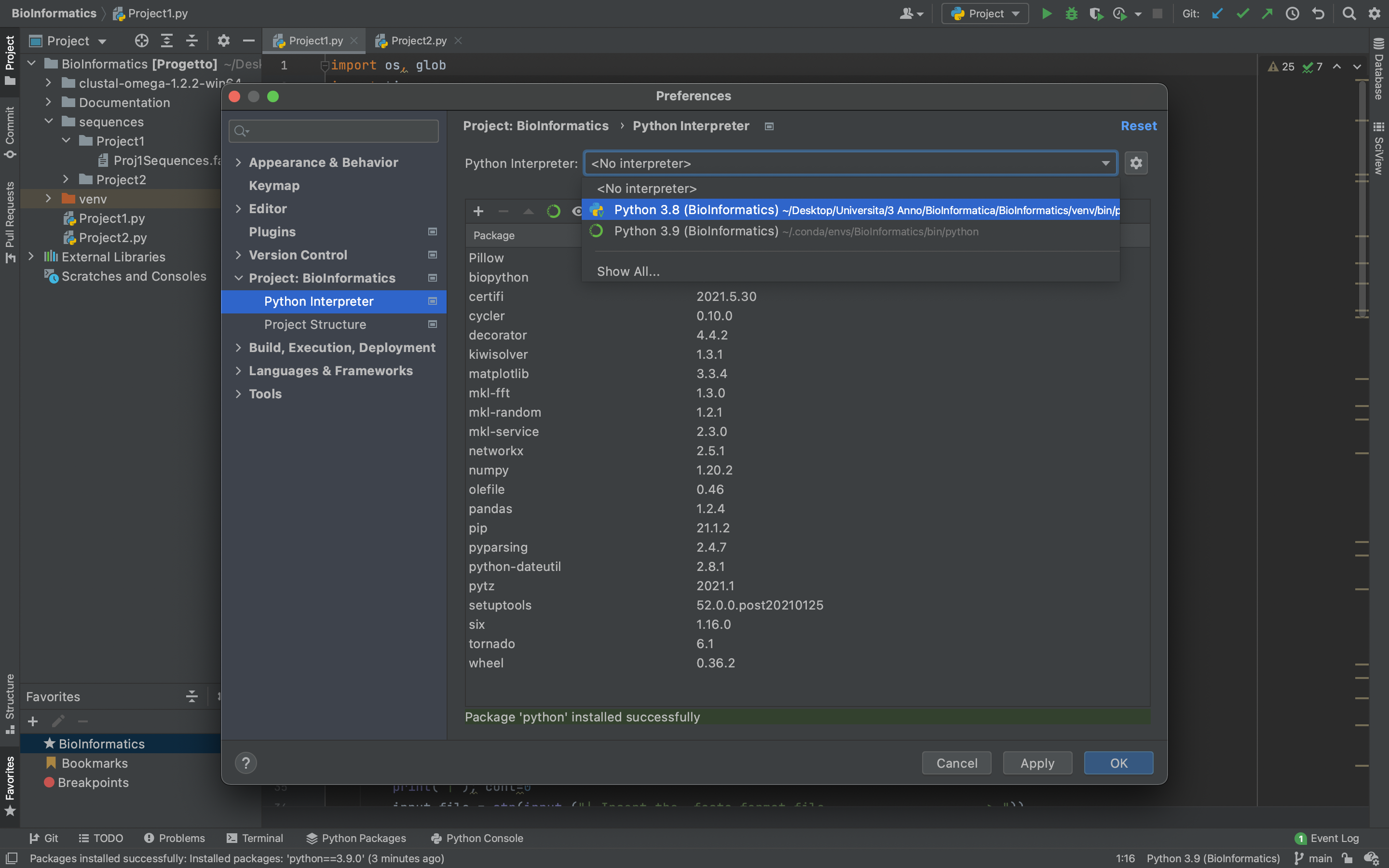
1. Configure PyCharm

Una volta installato l’IDE bisogna importare il “nome progetto” andando in: File > Open e selezionare la cartella contenente il software



Il prossimo passo da eseguire è configurare l’interprete del progetto.

Per fare ciò bisogna andare in: File > Settings selezionando la versione Python 3.8 (o maggiore).



Nel caso in cui non dovesse comparire, la procedura guidata dell’IDE permetterà di scaricarlo selezionando la freccia sottostante (venv) che permetterà di avere tutte le dipendenze installate



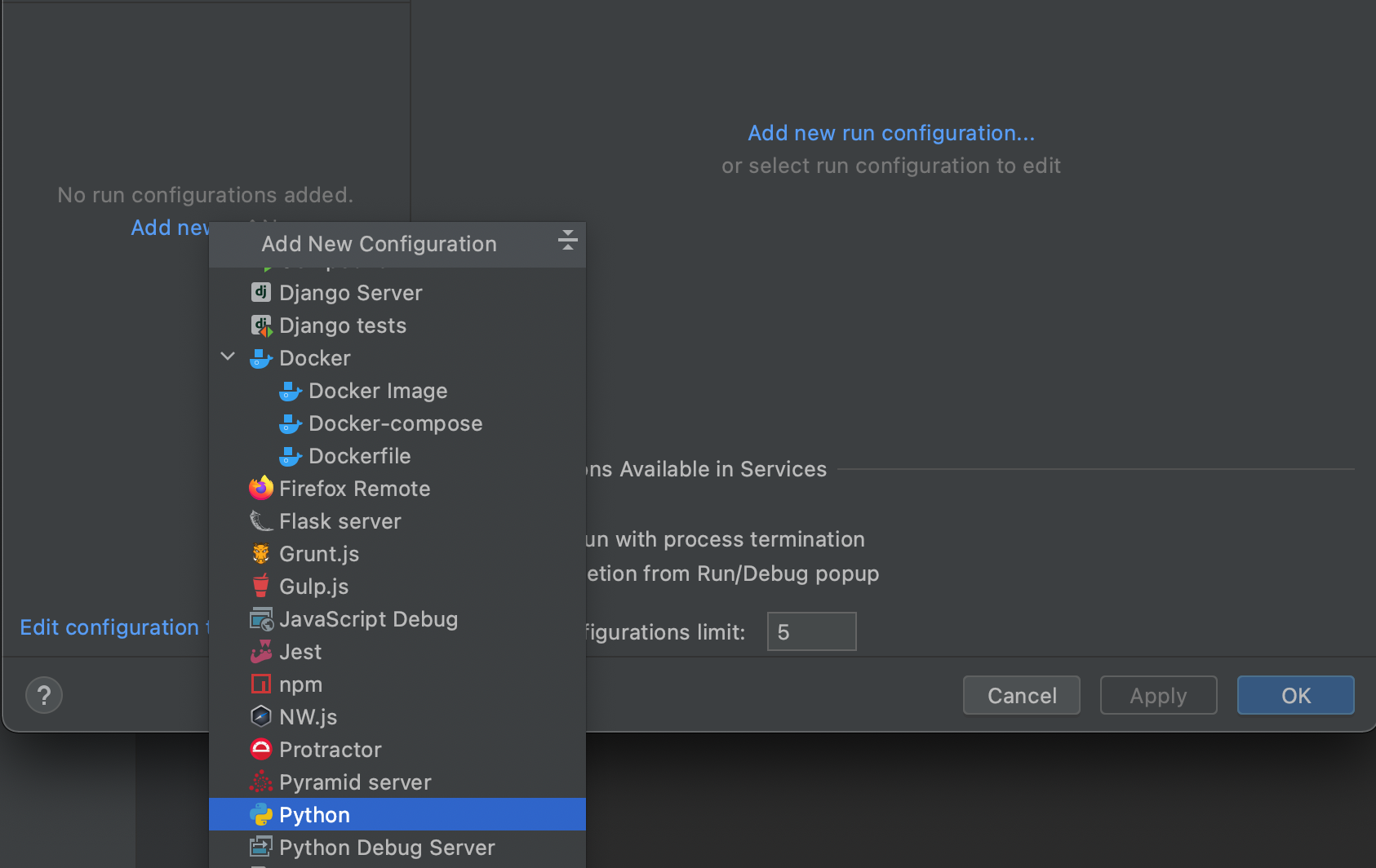
1. Create Configurations

Una volta importato il Progetto con il relativo interprete, bisognerà impostare la configurazione per eseguire il programma.

Per farlo, bisogna recarsi in alto a destra e cliccare su ‘Add Configuration’



E crearne una nuova

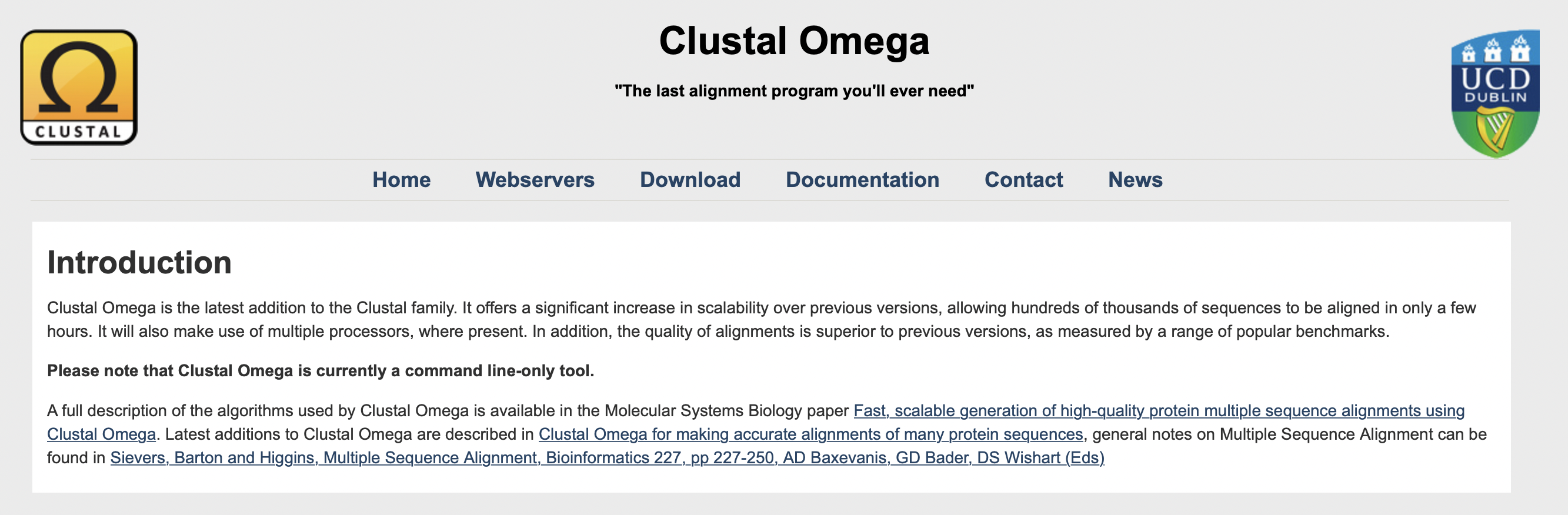


Una volta selezionato il path del file da eseguire e il path della cartella di lavoro, bisogna selezionare l’interprete di python configurato al passo 2 (precedente)

1. Clustal Omega

E’ un elemento fondamentale del software, che permetterà di eseguire un allineamento multiplo di un file contenente delle sequenze.

Per far funzionare il processo, bisognerà scaricarlo dal link.



A seconda del sistema utilizzato, si dovrà istallare il binario adatto.

Immagine che contiene testo

Descrizione generata automaticamente

1. Package (passo opzionale)

Per l’installazione dei package bisogna far uso del terminale che è già incorporato all’interno dell’IDE

Immagine che contiene testo, screenshot, monitor, portatile

Descrizione generata automaticamente

I package da installare esterni a python con i relativi comandi da eseguire sono:

* Biopython pip install biopython
* numpy pip install numpy
* terminaltables pip install terminaltables
* prettytable pip install prettytables
* matplotlib pip install matplotlib

Tutte le dipendenze mancanti, si identificano con un errore nella riga del comando di import, come in questo caso

Immagine che contiene testo, dispositivo, metro

Descrizione generata automaticamente

In questo modo inoltre, non solo si potrà capire se esiste questa libreria nel sistema, ma avremo anche un modo alternativo molto più user friendly per installare i package indispensabili

B. Running Program

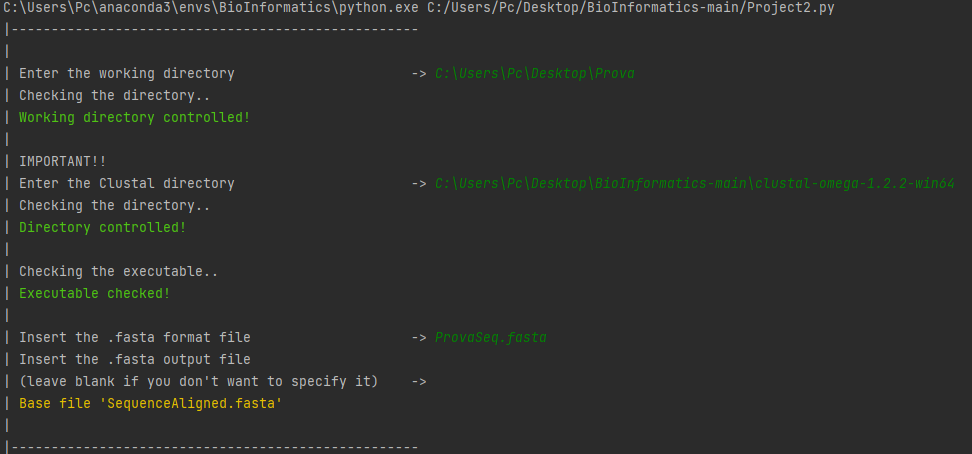
Una volta effettuati i passi mostrati sopra, bisognerà eseguire il software

Utilizzando la freccia nella barra in alto a destra



oppure Utilizzando i tasti rapidi (Maiusc + F10)

All’esecuzione del programma, verrà mostrato nella “python console” un menù che permetterà all’utente di inserire le corrette informazioni.



**PROJECT 1**

Questo programma, è l’implementazione di una pipeline che identifica le mutazioni tra le sequenze contenute all’interno di un file con estensione “.fasta”, restituendo in output una tabella contenente per ogni sequenza, le mutazioni della sequenza di riferimento identificata in ogni colonna tra le sequenze di riferimento

L’output sarà costituito in questo modo:

Immagine che contiene testo, elettronico, tastiera

Descrizione generata automaticamente

**PROJECT 2**

Il risultato di questo programma è l’implementazione di una pipeline che costruirà gli alberi filogenetici delle sequenze di 2 ceppi di Covid19.

Gli alberi che verranno costruiti sono

* Upgma
* Neighbor joining

In corrispondenza degli alberi prodotti, verrà stampata su schermo una tabella con le informazioni delle sequenze per agevolarne la lettura

