|  |  |
| --- | --- |
| **Practicum Case** |  |
| SCIE6062 | SCIE6062001 | SCIE6062016 | SCIE6062049  Computational Biology |
| **Mathematics & Statistics** | **E231-SCIE6062-VO01-05** |
| ***Valid on*** *Even Semester Year 2022/2023* | **Revision 00** |

**Learning Outcome**

* LO3 – implement basic bioinformatics analysis in scope of DNA composition and sequence analysis, protein synthesis, sequence alignment, and other works related to biological database using Biopython
* LO4 – analyze the basic bioinformatics analysis results using Biopython

## Topic

* Session 05 – Working with Biological GenBank Files

## Sub Topics

* Working with FASTA files
* Working with PDB files
* Protein Explanatory Data Analysis

## Materi

*Material*

1. How to download FASTA / GenBank file from the NCBI website:

* Go to the **NCBI website** at https://www.ncbi.nlm.nih.gov/.
* In the **search** **bar**, select “**Nucleotide**” from the drop-down options and type in the **name** of the sequence. Click on “**Search**”.
* Find the **desired entry** in the search results and click on its **title** to access its information page.
* On the information page, you can view the **details of the sequence**. To download the sequence data in **FASTA format**, click on the “**Send to**” button on the top right corner of the section. Then, select “**File**” as the destination type.
* Choose the **desired file format** (e.g., **FASTA**, **GenBank**) and click on the “**Create** **File**” button.
* The file will be **downloaded** to your computer, and you can then use it for your research.

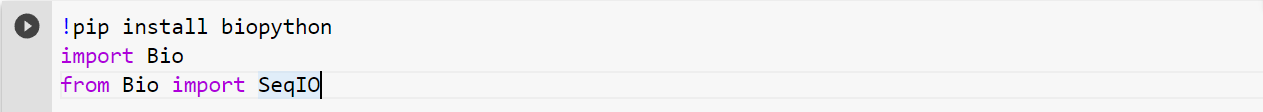
1. **Upload** the "**.fasta**" and "**.gb**" **files** to Google Colab's session storage by clicking the upload button or dragging and dropping the files.

Graphical user interface, text, application

Description automatically generated

1. **import Bio.SeqIO module**

The **Bio.SeqIO** module from the BioPython library is used for **reading and writing sequence data in a variety of formats**, including **fasta**, and **genbank**. It provides a unified interface for reading and writing sequence data, making it easy to switch between different file formats without having to change the code.



1. **Load** and **read** the “**.fasta**” file.

Graphical user interface, text, application

Description automatically generated

1. **Load** and **read** the “**.gb**” file.

