National
College of
Ireland

Configuration Manual

MSc Research Project
Data Analytics

Madhuri Dhatrak
Student ID: 21208808

School of Computing
National College of Ireland

Supervisor: Qurrat Ul Ain

National College of Ireland

MSc Project Submission Sheet

School of Computing

Student Name:	Madhuri Dhatrak ……
Student ID:	21208808 …….
Programme:	Master of Science in Data Analytics **Year:** SEP 2022-23. ……………………………………………
Module:	Research Project ………………………………………………………………………………………….………
Lecturer:	Qurrat Ul Ain ………………………………………………………………………………….………
Submission Due Date:	14/08/2023 ……………………………………………………………………………….………
Project Title:	Application of Supervised Learning Classifiers on Gut Microbial data to Predict Parkinson Disease ………………………………………………………………………….……
Word Count:	…………………………………… **Page Count:** ……………………………….………

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.
<u>ALL</u> internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	Madhuri Dhatrak ……………………………………………………………………………………………………
Date:	14/08/2023 ………………………………………………………………………………………………

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Configuration Manual

Madhuri Dhatrak
Student ID: 21208808

1   Introduction

The step-by-step procedure for setting up the environment and carrying out this research study is thoroughly described in this configuration handbook. This document contains details on the tools, software programming language, library packages, and system hardware settings.

The document also covers the steps like input, output artifacts and command line instructions required from installation step up to end outcomes of the research project.

2   Environment set up Instructions.

2.1   Software Specifications

The tools required for this project are as follows.

1) Based the Operating system environment QIIME can be installed on any platform as mentioned in install documentation (QIIME 2 user documentation, n.d.). This research project installed QIIME2 tool on Docker [1] .
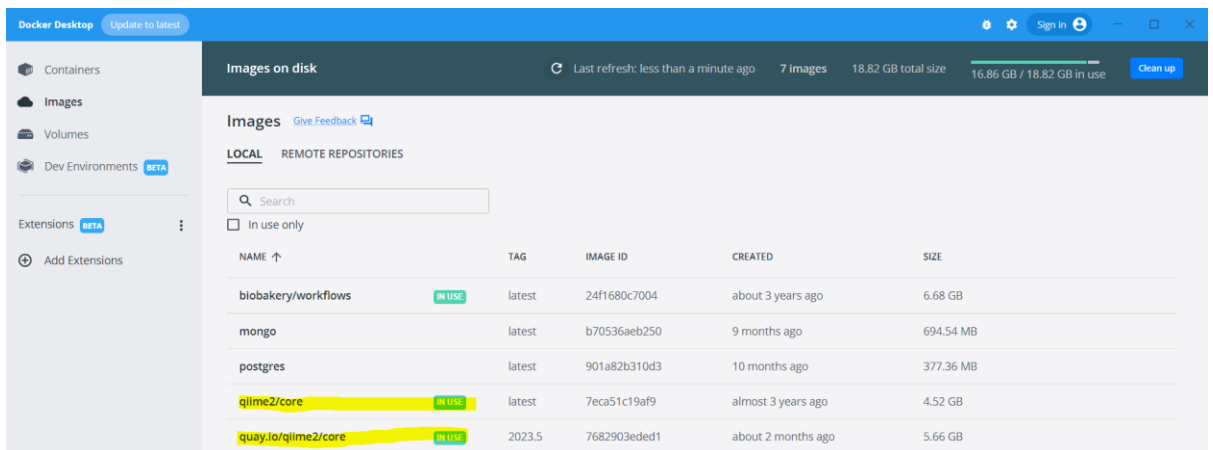
2) Latest Version QIIME 2 releases: 2023.5, (QIIME 2 user documentation, n.d.) is used for this project using Docker.This project installed QIIME on Docker .

---

[1] https://www.docker.com/products/docker-desktop/

Command Prompt - docker run -t -i -v C:/Users/madhu/:/data quay.io/qiime2/core:2023.5

```
Microsoft Windows [Version 10.0.19045.3208]
(c) Microsoft Corporation. All rights reserved.

C:\Users\madhu>docker run -t -i -v C:/Users/madhu/:/data quay.io/qiime2/core:2023.5
```

3) FastQC tool [2] for checking the quality of the sequences downloaded.

Software	Version
Docker	v0.11.2-desktop.1
QIIME 2	2023.5
FastQC	Version 0.12.0

2.2 Hardware Specifications

Below are the hardware specifications with which this research project is carried out.

Hardware	Specifications
Operating System	Windows 10 Home Version 22H2
Processor	Intel(R) Core (TM) i7-8550U CPU @ 1.80GHz 2.00 GHz
RAM	16.0 GB
Hard Disk	450 GB

2.3 Source data

The data for this is collected from open-source database European Nucleotide Archive (ENA) which is a repository storing annotated DNA and RNA sequences. The 17 selected samples from the project PRJEB27564 [3] are used for this study, out which 8 faecal samples are from healthy individuals and 9 faecal samples from Parkison's patient.

The raw sequence fastq files which are pair end reads with separate files for forward and reverse reads are first downloaded on to the local disk which is mounted on Docker Desktop.

---

[2] https://www.bioinformatics.babraham.ac.uk/projects/fastqc/
[3] https://www.ebi.ac.uk/ena/browser/view/PRJEB27564

```
Command Prompt - docker run -t -i -v C:/Users/madhu/:/data quay.io/qiime2/core:2023.5
Application config directory
/opt/conda/envs/qiime2-2023.5/var/q2cli

Getting help
To get help with QIIME 2, visit https://qiime2.org
(qiime2-2023.5) root@92247823ecf6:/data# wget -nc ftp.sra.ebi.ac.uk/vol1/run/ERR273/ERR2730240/P-53-O_R1_001.fastq.gz
Will not apply HSTS. The HSTS database must be a regular and non-world-writable file.
ERROR: could not open HSTS store at '/home/qiime2/.wget-hsts'. HSTS will be disabled.
--2023-08-12 10:33:55--  http://ftp.sra.ebi.ac.uk/vol1/run/ERR273/ERR2730240/P-53-O_R1_001.fastq.gz
Resolving ftp.sra.ebi.ac.uk (ftp.sra.ebi.ac.uk)... 193.62.193.138
Connecting to ftp.sra.ebi.ac.uk (ftp.sra.ebi.ac.uk)|193.62.193.138|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 392968 (384K) [application/x-gzip]
Saving to: 'P-53-O_R1_001.fastq.gz'

P-53-O_R1_001.fastq.gz        100%[===================================================================>] 383.76K  2.33MB/s    in 0.2s

2023-08-12 10:34:01 (2.33 MB/s) - 'P-53-O_R1_001.fastq.gz' saved [392968/392968]

(qiime2-2023.5) root@92247823ecf6:/data# wget -nc ftp.sra.ebi.ac.uk/vol1/run/ERR273/ERR2730240/P-53-O_R2_001.fastq.gz
Will not apply HSTS. The HSTS database must be a regular and non-world-writable file.
ERROR: could not open HSTS store at '/home/qiime2/.wget-hsts'. HSTS will be disabled.
--2023-08-12 10:34:05--  http://ftp.sra.ebi.ac.uk/vol1/run/ERR273/ERR2730240/P-53-O_R2_001.fastq.gz
Resolving ftp.sra.ebi.ac.uk (ftp.sra.ebi.ac.uk)... 193.62.193.138, 193.62.193.138
Connecting to ftp.sra.ebi.ac.uk (ftp.sra.ebi.ac.uk)|193.62.193.138|:80... connected.
HTTP request sent, awaiting response... 200 OK
Length: 468536 (458K) [application/x-gzip]
Saving to: 'P-53-O_R2_001.fastq.gz'

P-53-O_R2_001.fastq.gz        100%[===================================================================>] 457.55K   861KB/s    in 0.5s

2023-08-12 10:34:06 (861 KB/s) - 'P-53-O_R2_001.fastq.gz' saved [468536/468536]

(qiime2-2023.5) root@92247823ecf6:/data#
```

Below attachment gives the list of all the sample files used for this study downloaded from ENA repository.

Raw_seq_files_list_EN
A_download.txt

3   Implementation

Implementation for this study is mainly carried on QIIME 2 tool for the analysis of the 16SrRNA sequence to derive the features, ASV's and taxonomy abundance. With QIIME 2 installed on docker the implementation instructions are through command line interface.

All the commands from the beginning of installation to model evaluation is provided in the below attachment file.

q2cli_parkison_sampl
e_prediction.bash

3.1   Importing

The quality control of the downloaded raw files is first checked with external tool FastQC software.

Then the downloaded sample files on local disk are imported into the QIIME 2 tool for analysis. This tool as a framework to process these files (QIIME 2 development team, n.d.). Depending on the input files format and layout importing to QIIME 2 tool the steps differ. For this study the sample files are 16S rRNA demultiplexed pair end sequences fastq files. All possible file formats and importing instructions are provided in QIIME documentation (QIIME 2 user documentation, n.d.) (QIIME2 development team, n.d.).Manifest file format import is used,this file has sample id's and its absolute path for import.

```
sample-id    forward-absolute-filepath                                              reverse-absolute-filepath
C118    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-118-O_R1_001.fastq.gz    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-118-O_R2_001.fastq.gz
P60     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-60-O_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-60-O_R2_001.fastq.gz
C18     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-18-O_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-18-O_R2_001.fastq.gz
C30     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-30-O_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-30-O_R2_001.fastq.gz
C48     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-48-O_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-48-O_R2_001.fastq.gz
C51     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-51-O_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-51-O_R2_001.fastq.gz
C80     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-80-O_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-80-O_R2_001.fastq.gz
C82     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-82-O_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-82-O_R2_001.fastq.gz
C137    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-137-N_R1_001.fastq.gz    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/C-137-N_R2_001.fastq.gz
P16     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-16-N_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-16-N_R2_001.fastq.gz
P31     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-31-O_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-31-O_R2_001.fastq.gz
P53     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-53-O_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-53-O_R2_001.fastq.gz
P77     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-77-N_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-77-N_R2_001.fastq.gz
P99     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-99-N_R1_001.fastq.gz     /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-99-N_R2_001.fastq.gz
P103    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-103-N_R1_001.fastq.gz    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-103-N_R2_001.fastq.gz
P116    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-116-O_R1_001.fastq.gz    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-116-O_R2_001.fastq.gz
P120    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-120-O_R1_001.fastq.gz    /data/ENA_PRJEB27564_SAMPLES/ENA_PRJEB27564_SAMPLES/P-120-O_R2_001.fastq.gz
```

Command Prompt - docker run -t -i -v C:/Users/madhu/:/data quay.io/qiime2/core:2023.5

```
(qiime2-2023.5) root@92247823ecf6:/data#
(qiime2-2023.5) root@92247823ecf6:/data# qiime tools import \
>    --type 'SampleData[PairedEndSequencesWithQuality]' \
>    --input-format PairedEndFastqManifestPhred33V2 \
>    --input-path ./pair_end_manifest.tsv \
>    --output-path ./paired-end-demux.qza
Imported ./pair_end_manifest.tsv as PairedEndFastqManifestPhred33V2 to ./paired-end-demux.qza
(qiime2-2023.5) root@92247823ecf6:/data# qiime demux summarize \
>    --i-data ./paired-end-demux.qza \
>    --o-visualization ./paired-end-demux_seqs.qzv
Saved Visualization to: ./paired-end-demux_seqs.qzv
(qiime2-2023.5) root@92247823ecf6:/data#
```

QIIME2 offers a interface to the view .qza and .qzv artifacts files on a brower directly. (Anon., n.d.).
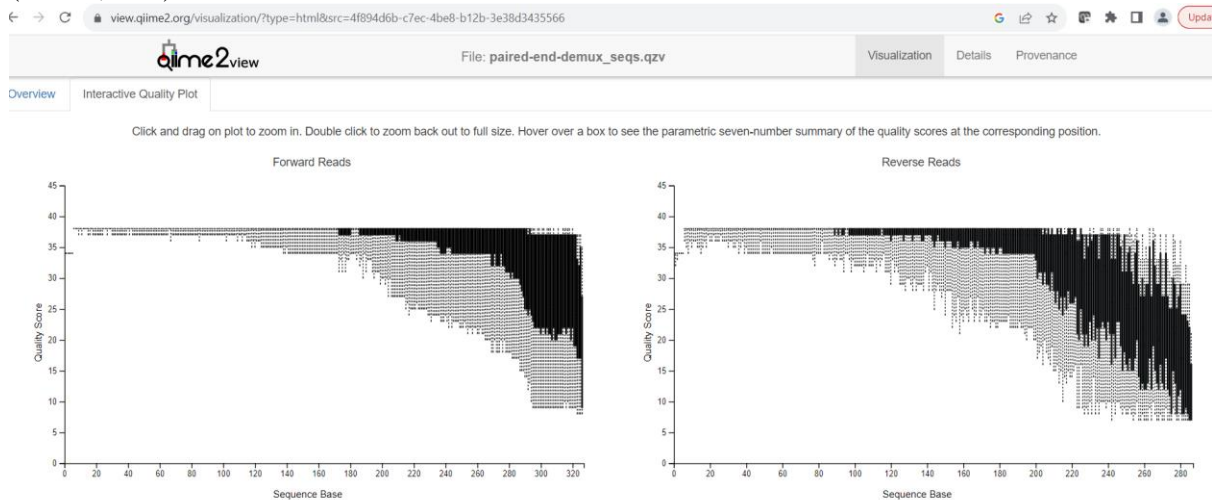
3.2 Data Analysis

In this step sequence quality control is carried out. Denoising the sequences using deblur methods and feature table and representative sequences are generated. This is the most important and critical step for further analysis as the artifacts generated at this step are used for various analysis like taxonomy classification, finding alpha diversity, beta diversity and sample classification.

A) Joining Pair end Reads:

As the fastq files are pair end reads, sequences need to be joined because deblur method works on single end layout files. Using merge pairs in vsearch method forward and reverse read files are joined. Below statements accomplish this task in QIIME 2.


```
Select Command Prompt - docker run -t -i -v C:/Users/madhu/:/data quay.io/qiime2/core:2023.5
(qiime2-2023.5) root@92247823ecf6:/data# Y
bash: Y: command not found
(qiime2-2023.5) root@92247823ecf6:/data#   qiime vsearch merge-pairs \
>       --i-demultiplexed-seqs paired-end-demux.qza \
>       --o-merged-sequences paired-joined-reads.qza \
>       --p-threads 8
Saved SampleData[JoinedSequencesWithQuality] to: paired-joined-reads.qza
(qiime2-2023.5) root@92247823ecf6:/data# qiime quality-filter q-score \
>       --i-demux paired-joined-reads.qza \
>       --o-filtered-sequences paired-joined-filtered.qza \
>       --o-filter-stats paired-joined-filter-stats.qza
Saved SampleData[JoinedSequencesWithQuality] to: paired-joined-filtered.qza
Saved QualityFilterStats to: paired-joined-filter-stats.qza
(qiime2-2023.5) root@92247823ecf6:/data# qiime demux summarize \
>   --i-data paired-joined-filtered.qza \
>   --o-visualization paired-joined-filtered.qzv
Saved Visualization to: paired-joined-filtered.qzv
```

B) Denoising using Deblur:

Once the pair end read are joined the out artifacts are single end reads and these reads are input artifacts to the deblur denoise-16S method available in QIIME 2 as a method. The sequences are trimmed to length 300 where quality score of the sequences started to decrease.

Feature table view

Denoise statistics view

Representative sequence's view

C) Taxonomy Classification using Feature Classifier:

Once the feature table is extracted by the deblur method, further analysis on taxonomy classification using pretrained classifier with silva and green genes reference database is performed.

Taxonomy summary report is generated to see the levels and abundance of the microbial species per feature. With level of confident the species is predicted with reference database.

```
Select Command Prompt - docker  run -t -i -v C:/Users/madhu/:/data quay.io/qiime2/core:2023.5

(qiime2-2023.5) root@92247823ecf6:/data# qiime taxa collapse    --i-table paired-table-deblur.qza
silva.qza
Saved FeatureTable[Frequency] to: paired-collapsed-deblur-table_silva.qza
(qiime2-2023.5) root@92247823ecf6:/data#  qiime taxa collapse \
>    --i-table paired-table-deblur.qza \
>    --i-taxonomy paired-deblur-taxonomy_gg.qza \
>    --p-level 6 \
>    --o-collapsed-table paired-collapsed-deblur-table_gg.qza
Saved FeatureTable[Frequency] to: paired-collapsed-deblur-table_gg.qza
(qiime2-2023.5) root@92247823ecf6:/data# qiime feature-table relative-frequency \
>    --i-table paired-collapsed-deblur-table_gg.qza \
>    --o-relative-frequency-table paired-relative-abundance-deblur_gg.qza
Saved FeatureTable[RelativeFrequency] to: paired-relative-abundance-deblur_gg.qza
(qiime2-2023.5) root@92247823ecf6:/data# qiime feature-table summarize \
>    --i-table paired-relative-abundance-deblur_gg.qza \
>    --o-visualization paired-relative-abundance-deblur-summary_gg.qzv

Saved Visualization to: paired-relative-abundance-deblur-summary_gg.qzv
(qiime2-2023.5) root@92247823ecf6:/data#
(qiime2-2023.5) root@92247823ecf6:/data#   qiime feature-table filter-samples \
>    --i-table ./paired-table-deblur.qza \
>    --p-min-frequency 2000 \
>    --o-filtered-table ./table_2k.qza
Saved FeatureTable[Frequency] to: ./table_2k.qza
(qiime2-2023.5) root@92247823ecf6:/data#  qiime taxa barplot \
>    --i-table ./table_2k.qza \
>    --i-taxonomy ./paired-deblur-taxonomy_silva.qza \
>    --m-metadata-file ./pair_end_meta.tsv \
>    --o-visualization ./taxa_barplot_silva.qzv
Saved Visualization to: ./taxa_barplot_silva.qzv
(qiime2-2023.5) root@92247823ecf6:/data#  qiime taxa barplot \
>    --i-table ./table_2k.qza \
>    --i-taxonomy ./paired-deblur-taxonomy_gg.qza \
>    --m-metadata-file ./pair_end_meta.tsv \
>    --o-visualization ./taxa_barplot_gg.qzv
Saved Visualization to: ./taxa_barplot_gg.qzv
```

3.3 Model

Supervised machine learning classifier algorithms built based on sci-kit learn python packages provided in QIIME 2 to predict the samples susceptible to Parkinson's disease. —p -estimator argument to the plugin gives a list of classifiers to choose (scikit-learn developers, n.d.). Meta data file is provided as one of the input artifacts to the sample classifier which give the samples meta data as shown. Based the meta data column DieseaseStatus which serves as labelled column for the model training and predicts the test samples.

SampleID	Group	Description	SubjectID	TimePoint		Hostbodyproduct	Geographiclocation	Age	Sex	DiseaseStatus
C118	Control	human-gut	C-118-O	baseline	Stool	Finland	N/A	N/A	HC	
P60	Case	human-gut	P-60-O	baseline	Stool	Finland	N/A	N/A	PD	
P116	Case	human-gut	P-116-O	baseline	Stool	Finland	N/A	N/A	PD	
P120	Case	human-gut	P-120-O	baseline	Stool	Finland	N/A	N/A	PD	
C80	Control	human-gut	C-80-O	baseline	Stool	Finland	N/A	N/A	HC	
C51	Control	human-gut	C-51-O	baseline	Stool	Finland	N/A	N/A	HC	
C82	Control	human-gut	C-82-O	baseline	Stool	Finland	N/A	N/A	HC	
P77	Case	human-gut	P-77-O	baseline	Stool	Finland	N/A	N/A	PD	
C18	Control	human-gut	C-18-O	baseline	Stool	Finland	N/A	N/A	HC	
C48	Control	human-gut	C-48-O	baseline	Stool	Finland	N/A	N/A	HC	
P103	Case	human-gut	P-103-O	baseline	Stool	Finland	N/A	N/A	PD	
C137	Control	human-gut	C-137-O	baseline	Stool	Finland	N/A	N/A	HC	
C30	Control	human-gut	C-30-O	baseline	Stool	Finland	N/A	N/A	HC	
P99	Case	human-gut	P-99-O	baseline	Stool	Finland	N/A	N/A	PD	
P16	Case	human-gut	P-16-O	baseline	Stool	Finland	N/A	N/A	PD	
P31	Case	human-gut	P-31-O	baseline	Stool	Finland	N/A	N/A	PD	
P53	Case	human-gut	P-53-O	baseline	Stool	Finland	N/A	N/A	PD	

Sample classifier with default inputs

Command Prompt - docker run -t -i -v C:/Users/madhu/:/data quay.io/qiime2/core:2023.5

```
(qiime2-2023.5) root@92247823ecf6:/data# qiime sample-classifier classify-samples \
>    --i-table ./paired-table-deblur.qza \
>    --m-metadata-file ./pair_end_meta.tsv\
>    --m-metadata-column DiseaseStatus \
>    --output-dir ./pairend-classifier-results-E3
Saved SampleEstimator[Classifier] to: ./pairend-classifier-results-E3/sample_estimator.qza
Saved FeatureData[Importance] to: ./pairend-classifier-results-E3/feature_importance.qza
Saved SampleData[ClassifierPredictions] to: ./pairend-classifier-results-E3/predictions.qza
Saved Visualization to: ./pairend-classifier-results-E3/model_summary.qzv
Saved Visualization to: ./pairend-classifier-results-E3/accuracy_results.qzv
Saved SampleData[Probabilities] to: ./pairend-classifier-results-E3/probabilities.qza
Saved Visualization to: ./pairend-classifier-results-E3/heatmap.qzv
Saved SampleData[TrueTargets] to: ./pairend-classifier-results-E3/training_targets.qza
Saved SampleData[TrueTargets] to: ./pairend-classifier-results-E3/test_targets.qza
```

Sample classifier with estimator Random Forest Classifier with n estimators decision tress to 50.

Command Prompt - docker run -t -i -v C:/Users/madhu/:/data quay.io/qiime2/core:2023.5

```
(qiime2-2023.5) root@92247823ecf6:/data# qiime sample-classifier classify-samples \
>    --i-table paired-table-deblur.qza \
>    --m-metadata-file pair_end_meta.tsv \
>    --m-metadata-column DiseaseStatus \
>    --p-optimize-feature-selection \
>    --p-parameter-tuning \
>    --p-estimator RandomForestClassifier \
>    --p-n-estimators 50 \
>    --p-test-size 0.3 \
>    --output-dir paired-end-classifier-E4
Saved SampleEstimator[Classifier] to: paired-end-classifier-E4/sample_estimator.qza
Saved FeatureData[Importance] to: paired-end-classifier-E4/feature_importance.qza
Saved SampleData[ClassifierPredictions] to: paired-end-classifier-E4/predictions.qza
Saved Visualization to: paired-end-classifier-E4/model_summary.qzv
Saved Visualization to: paired-end-classifier-E4/accuracy_results.qzv
Saved SampleData[Probabilities] to: paired-end-classifier-E4/probabilities.qza
Saved Visualization to: paired-end-classifier-E4/heatmap.qzv
Saved SampleData[TrueTargets] to: paired-end-classifier-E4/training_targets.qza
Saved SampleData[TrueTargets] to: paired-end-classifier-E4/test_targets.qza
```

Sample classifier with estimator LinearSVC


```
(qiime2-2023.5) root@92247823ecf6:/data#   qiime sample-classifier classify-samples \
>   --i-table paired-table-deblur.qza \
>   --m-metadata-file pair_end_meta.tsv \
>   --m-metadata-column DiseaseStatus \
>   --p-optimize-feature-selection \
>   --p-parameter-tuning \
>   --p-test-size 0.3 \
>   --p-estimator LinearSVC \
>   --p-random-state 123 \
>   --output-dir paired-end-classifier-E5
Saved SampleEstimator[Classifier] to: paired-end-classifier-E5/sample_estimator.qza
Saved FeatureData[Importance] to: paired-end-classifier-E5/feature_importance.qza
Saved SampleData[ClassifierPredictions] to: paired-end-classifier-E5/predictions.qza
Saved Visualization to: paired-end-classifier-E5/model_summary.qzv
Saved Visualization to: paired-end-classifier-E5/accuracy_results.qzv
Saved SampleData[Probabilities] to: paired-end-classifier-E5/probabilities.qza
Saved Visualization to: paired-end-classifier-E5/heatmap.qzv
Saved SampleData[TrueTargets] to: paired-end-classifier-E5/training_targets.qza
Saved SampleData[TrueTargets] to: paired-end-classifier-E5/test_targets.qza
```

Your third section. Change the header and label to something appropriate.

Model parameters

	Parameter setting
est__bootstrap	True
est__ccp_alpha	0.0
est__class_weight	None
est__criterion	entropy
est__max_depth	16
est__max_features	0.1
est__max_leaf_nodes	None
est__max_samples	None

4   References

Anon., n.d. *QIIME2 View.* [Online]
Available at: https://view.qiime2.org/
QIIME 2 development team, n.d. *Overview of QIIME 2 Plugin Workflows.* [Online]
Available at: https://docs.qiime2.org/2023.5/tutorials/overview/#denoising-and-clustering
QIIME 2 user documentation, n.d. *Installing QIIME 2.* [Online]
Available at: https://docs.qiime2.org/2023.5/install/
[Accessed 2023].

QIIME2 development team, n.d. *Importing data.* [Online]
Available at: https://docs.qiime2.org/2023.5/tutorials/importing/
scikit-learn developers, n.d. *scikit-learn.* [Online]
Available at: https://scikit-learn.org/stable/supervised_learning