ECOLE POLYTECHNIQUE FEDERALE DE LAUSANNE
SCHOOL OF LIFE SCIENCES

Master project in Life Sciences Engineering

[THESIS TITLE]

Carried out in the laboratory of [SUPERVISOR NAME]
at [INSTITUTION/LABORATORY]
Under the supervision of [CO-SUPERVISOR NAME]

Done by

[STUDENT NAME]

Under the direction of
[EPFL SUPERVISOR NAME]

EPFL

LAUSANNE, EPFL 2025

Acknowledgements

I thank [EXTERNAL SUPERVISOR NAME], [EPFL SUPERVISOR NAME], and [DAILY SUPERVISOR/POSTDOC NAME] for their guidance and support, my lab colleagues for a collaborative environment, and my family and friends for their encouragement throughout this work.

Abstract

This thesis addresses a broadly relevant problem in life sciences engineering and proposes a concise, reproducible, and data-driven approach for analyzing complex biological data. The objectives are to standardize the workflow, ensure transparent evaluation, and support practical use across related contexts.
We present a general pipeline that integrates preprocessing, statistical modeling, and machine learning, accompanied by clear reporting and open templates. The results highlight consistent patterns across datasets (without task-specific claims), and the contributions include a reusable framework, documented protocols, and guidance for future applications and extensions.

Acronyms

Acronym	Definition
AI	Artificial Intelligence
API	Application Programming Interface
AUC	Area Under the Curve
CV	Cross Validation
DNA	Deoxyribonucleic Acid
EPFL	École Polytechnique Fédérale de Lausanne
FDR	False Discovery Rate
GPU	Graphics Processing Unit
HPC	High-Performance Computing
ML	Machine Learning
mRNA	messenger Ribonucleic Acid
PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
QC	Quality Control
RF	Random Forest
RNA	Ribonucleic Acid
ROC	Receiver Operating Characteristic
SLS	School of Life Sciences
SVM	Support Vector Machine
UI	User Interface

Contents

1 Introduction

This template provides a brief, generic introduction suitable for an EPFL School of Life Sciences master thesis.

1.1 Figures

Figure 1 shows how to insert and reference a single image. In Figure 2, panels a and b demonstrate subfigures and cross referencing to specific panels.

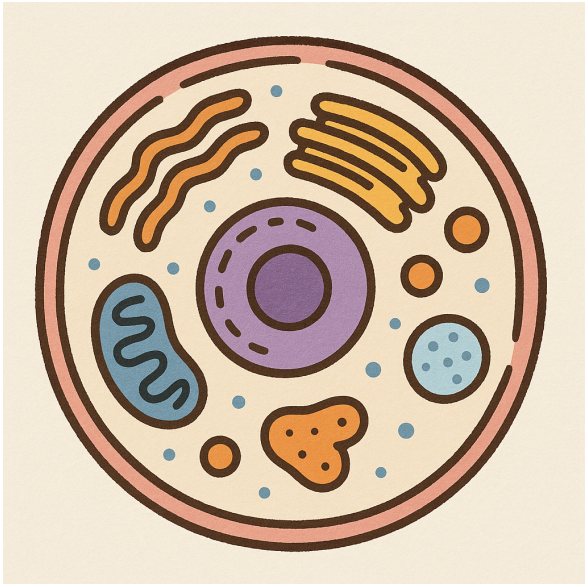
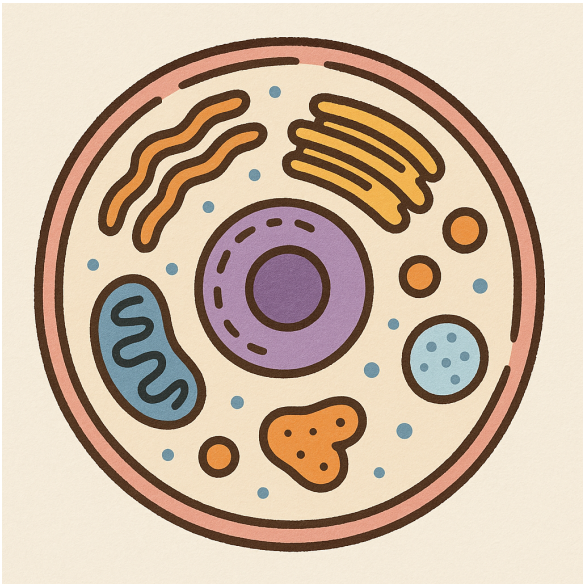
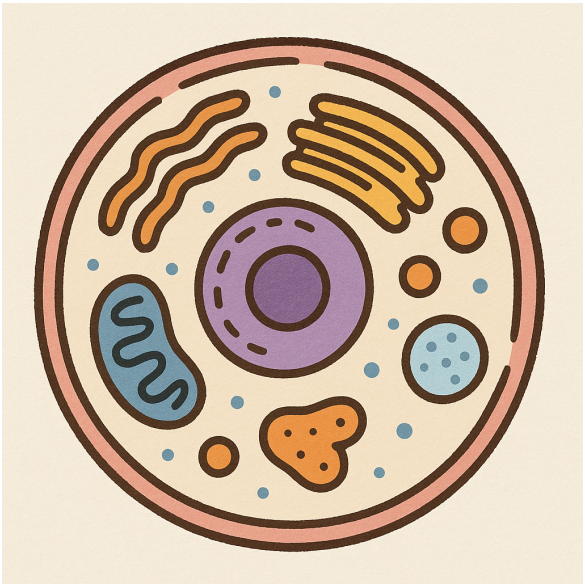


Figure 1: Overview illustration: a mother cell giving rise to two daughter cells.



(a) Daughter Cell 1.



(b) Daughter Cell 2.

Figure 2: Subfigure example: the same image reused to illustrate two conceptual panels.

2 Materials and Methods

This section outlines a concise, generic workflow suitable for life sciences data analysis. It typically includes data collection or selection, preprocessing and quality control, feature engineering, model development, and evaluation under reproducible settings.
Replace this placeholder with project specific details: datasets and inclusion criteria, preprocessing steps, models and hyperparameters, validation procedures, software and versions, and any ethical or data use considerations.

2.1 Example of Citation Method

Reproducible computational research and rigorous data stewardship are foundational to modern life sciences. Adopting the FAIR Guiding Principles ensures that datasets and metadata are findable, accessible, interoperable, and reusable across studies and platforms [1]. Following established best practices for reproducibility—such as version control, scripted analyses, exact environment capture, and public archiving—reduces analytic ambiguity and supports transparent validation [2]. In parallel, advances in machine learning provide scalable tools for pattern discovery and prediction from high-dimensional biological measurements, but their utility depends critically on well-curated, shareable data and reproducible workflows [3].

3 Results

This section summarizes representative outcomes at a high level, focusing on patterns aligned with the study objectives and on robustness checks (e.g., cross validation, external validation, and uncertainty estimates).
Replace this placeholder with project specific findings: key metrics and confidence intervals, brief references to figures/tables, and observations relevant to the research questions. Avoid over claiming; note caveats that affect interpretation and generalizability.

3.1 Tables

Table 1 illustrates how to import a table from a CSV file (see `tables/results_example.csv`). Update the CSV and recompile to refresh the table.

Metric	A	B
Accuracy	0.81	0.78
Precision	0.79	0.75
Recall	0.80	0.76
F1-score	0.795	0.755
AUC	0.84	0.82

Table 1: Dummy performance metrics imported from CSV.

Large tables can overflow the page. Table 2 shows an example of a table that is too wide. Table 3 shows how to wrap a wide table with :

```
\resizebox{\textwidth}{!}{...}
```

so it fits the page width.

ID	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
A01	0.12	0.34	0.56	0.78	0.91	0.45	0.67	0.89	0.23	0.35	0.44	0.52	0.68	0.72	0.81
A02	0.22	0.31	0.58	0.74	0.88	0.41	0.63	0.85	0.27	0.39	0.48	0.54	0.70	0.74	0.83
A03	0.18	0.29	0.61	0.72	0.86	0.47	0.66	0.83	0.25	0.37	0.46	0.50	0.66	0.71	0.80
A04	0.20	0.33	0.59	0.76	0.90	0.43	0.65	0.87	0.21	0.31	0.42	0.49	0.64	0.69	0.78

Table 2: Example of a wide table resized to fit the page width.

ID	F1	F2	F3	F4	F5	F6	F7	F8	F9	F10	F11	F12	F13	F14	F15
A01	0.12	0.34	0.56	0.78	0.91	0.45	0.67	0.89	0.23	0.35	0.44	0.52	0.68	0.72	0.81
A02	0.22	0.31	0.58	0.74	0.88	0.41	0.63	0.85	0.27	0.39	0.48	0.54	0.70	0.74	0.83
A03	0.18	0.29	0.61	0.72	0.86	0.47	0.66	0.83	0.25	0.37	0.46	0.50	0.66	0.71	0.80
A04	0.20	0.33	0.59	0.76	0.90	0.43	0.65	0.87	0.21	0.31	0.42	0.49	0.64	0.69	0.78

Table 3: Example of a wide table resized to fit the page width.

4 Discussion

This section provides a concise interpretation of the results in relation to the study objectives. Emphasis is placed on what the findings suggest, how robust they appear under the chosen validation strategy, and how they fit within a broader life sciences context.
Replace this placeholder with project specific discussion points: main takeaways, practical implications, limitations (data, methods, assumptions), and prioritized future directions. Clearly distinguish evidence based claims from speculation, and note any dependencies that may affect generalizability.

5 Conclusion

This section provides a brief synthesis of the work and its contributions. In its generic form, it highlights that a reproducible, data driven methodology was presented, representative results were obtained, and practical lessons were drawn for life sciences applications.
Replace this placeholder with project specific takeaways: summarize the main findings without restating all details, note limitations that matter most, and outline concrete next steps (e.g., additional validation, broader datasets, or methodological extensions).

6 Supplementary Material

This section provides space for additional material that supports the main text, such as extended methods, extra figures/tables, or robustness checks. Include only information that aids reproducibility or clarifies key results.
Replace this placeholder with project specific supplementary items (e.g., detailed parameter settings, expanded validation, or additional results). Reference supplementary items from the main text where appropriate.

References

1. Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3** (2016).
2. Sandve, G. K., Nekrutenko, A., Taylor, J. & Hovig, E. Ten Simple Rules for Reproducible Computational Research. *PLOS Computational Biology* **9,** e1003285 (2013).
3. Tarca, A. L., Carey, V. J., Chen, X.-w., Romero, R. & Drăghici, S. Machine Learning and Its Applications to Biology. *PLOS Computational Biology* **3,** e116 (2007).