

# Contents

<b>1</b>	<b>Introduction heading 1</b>	<b>1</b>
1.1	Labelled Heading 2 . . . . .	1
1.2	A 3rd heading 2 . . . . .	1
1.3	<i>Arabidopsis thaliana</i> . . . . .	2
<b>2</b>	<b>This ones a list set or something</b>	<b>2</b>
<b>3</b>	<b>Now for a table</b>	<b>3</b>
<b>A</b>	<b>Gene Lists</b>	<b>3</b>
A.1	<i>Arabidopsis thaliana</i> . . . . .	3
<b>B</b>	<b>Matlab Scripts</b>	<b>3</b>
B.1	MLP Class . . . . .	4
	<b>References</b>	<b>4</b>

## 1 Introduction heading 1

A basic template. For example,

Bioinformatics is interpreting signals from biological systems as information to understand aspects of the system [1].

### 1.1 Labelled Heading 2

A section with a label in the heading.

### 1.2 A 3rd heading 2

Linking back to the previous section 1.1.

Despite the broad variety of gene prediction software systems . . . as their published performance using plant data is almost always restricted to the model plant *Arabidopsis thaliana*[2].

### 1.3 *Arabidopsis thaliana*

*Arabidopsis thaliana* (*A. thaliana*), commonly water-cress, is the most extensively used model plant organism in plant biology research, largely due to the early adoption of the species as a model organism in the early 1980s[3].



Figure 1: *A. thaliana*

Unlike the plant in figure 1, next we have an equation 1. Note, the  $\alpha$  parameter controls the transition gradient of the output curve

$$f(x) = \frac{1}{(1 + e^{-\alpha x})} \quad (1)$$

## 2 This ones a list set or something

Listing 1: A single RNA-Seq read

```
@SRR9718417.6 6 length=150
NGCTGCCTGATGGGCCCTTTGGCTTGACAACAGTCTTCCTCATTGTCACACGGCC ...
+SRR9718417.6 6 length=150
#AAFFKKKKKKKAKKKK<KKKKKKKKFKKKKKKKKKKKKKKKKKKKKKKKKKKKKK ...
```

### 3 Now for a table

The top 10 performing models from the *A. thaliana* experimental data validation results for all high-complexity genes are shown in table 1, displaying the models' metrics, and

Model	Accuracy	STD	Sensitivity	STD	Specificity	STD	Inf.	STD
1	0.945	0.131	0.918	0.228	0.981	0.042	0.899	0.238
2	0.945	0.131	0.920	0.228	0.980	0.042	0.899	0.238
3	0.945	0.131	0.918	0.228	0.981	0.042	0.899	0.238
4	0.945	0.131	0.918	0.228	0.981	0.042	0.899	0.238
5	0.945	0.131	0.918	0.228	0.981	0.042	0.899	0.238
6	0.945	0.131	0.918	0.228	0.981	0.042	0.899	0.238
7	0.945	0.131	0.918	0.228	0.980	0.042	0.899	0.238
8	0.945	0.131	0.918	0.228	0.981	0.042	0.899	0.238
9	0.944	0.131	0.918	0.228	0.981	0.042	0.899	0.238
10	0.944	0.131	0.917	0.228	0.981	0.042	0.899	0.238

Table 1: Top 10 model metrics - high-complexity

## A Gene Lists

### A.1 *Arabidopsis thaliana*

Heres 2

Table 2: *Actinidia chinensis* gene lists

<i>A. thaliana</i> Training	<i>A. thaliana</i> Simulated	<i>A. thaliana</i> Validation
AT1G01320	AT1G55860	AT3G48190
AT1G02080	AT1G13280	AT3G50380
AT1G03060	AT1G07480	AT3G50590
AT1G05570	AT1G22140	AT3G59100

## B Matlab Scripts

## B.1 MLP Class

```
% MLP
%
% 1+ hidden-layer perceptron model.
% Can handle as many hidden layers as needed.
%
% Usage:
%
% Create a MLP with the constructor:
%
% myMLP = MLP([layer-sizes], learning-rate, bias, activation-function,
%              alpha)
```

## References

Control bib title

- [1] J. Pevsner, *Bioinformatics and functional genomics*. John Wiley & Sons, 2015.
- [2] H. Xie, B. Dain, D. R. Becker, and R. E. Drake, “Current methods of gene prediction, their strengths and weaknesses,” *Rehabilitation Counseling Bulletin*, vol. 40, no. 4, pp. 230–239, 1997, doi: 10.1093/nar/gkf543.
- [3] D. W. Meinke, J. M. Cherry, C. Dean, S. D. Rounsley, and M. Koornneef, “*Arabidopsis thaliana*: a model plant for genome analysis,” *Science*, vol. 282, no. 5389, pp. 662–682, 1998.