## MogaML

Generated by Doxygen 1.9.6

I File Index	1
1.1 File List	1
2 File Documentation	3
2.1 src/dimreduction.c File Reference	3
2.1.1 Function Documentation	3
2.1.1.1 ChiSquared_Reduce()	3
2.1.1.2 PCA_Reduce()	4
2.2 src/main.c File Reference	4
2.2.1 Function Documentation	4
2.2.1.1 main()	4
2.3 src/matrix.c File Reference	4
2.3.1 Macro Definition Documentation	6
2.3.1.1 ERROR	6
2.3.2 Function Documentation	6
2.3.2.1 Matrix_add()	6
2.3.2.2 Matrix_approx_equal()	7
2.3.2.3 Matrix_clone()	7
2.3.2.4 Matrix_col()	8
2.3.2.5 Matrix_det()	8
2.3.2.6 Matrix_display()	9
2.3.2.7 Matrix_equal()	9
2.3.2.8 Matrix_free()	9
2.3.2.9 Matrix_frobenius_norm()	10
2.3.2.10 Matrix_from_array()	10
2.3.2.11 Matrix_identity()	10
2.3.2.12 Matrix_inverse()	11
2.3.2.13 Matrix_minor()	11
2.3.2.14 Matrix_multiply()	12
2.3.2.15 Matrix_norm()	12
2.3.2.16 Matrix_reset()	13
2.3.2.17 Matrix_row()	13
2.3.2.18 Matrix_scale()	13
2.3.2.19 Matrix_slice_rows()	14
2.3.2.20 Matrix_solve()	14
2.3.2.21 Matrix_sub()	15
2.3.2.22 Matrix_submatrix()	15
2.3.2.23 Matrix_svd()	16
2.3.2.24 Matrix_trace()	16
2.3.2.25 Matrix_transpose()	17
2.3.2.26 Matrix_zeros()	17
2.3.2.27 Vector_max()	17

2.3.2.28 Vector_max_index()	18
2.3.2.29 Vector_norm()	18
2.3.3 Variable Documentation	19
2.3.3.1 MATRIX_MAX_ITER	19
2.3.3.2 MATRIX_TOLERANCE	19
2.4 src/neuralnetwork.c File Reference	19
2.4.1 Function Documentation	19
2.4.1.1 NN_backward_pass()	19
2.4.1.2 NN_create_layer()	20
2.4.1.3 NN_create_network()	20
2.4.1.4 NN_create_neuron()	20
2.4.1.5 NN_forward_pass()	20
2.4.1.6 NN_get_sigmoid()	20
2.5 src/supervised.c File Reference	20
2.5.1 Macro Definition Documentation	22
2.5.1.1 ERROR	22
2.5.2 Function Documentation	22
2.5.2.1 GaussianNBC()	22
2.5.2.2 GaussianNBC_predict()	23
2.5.2.3 KNN_append_data()	23
2.5.2.4 KNN_free()	24
2.5.2.5 KNN_predict()	24
2.5.2.6 KNNClassifier()	24
2.5.2.7 KNNRegressor()	25
2.5.2.8 LassoRegression()	25
2.5.2.9 LassoRegression_compute_gradient()	26
2.5.2.10 LassoRegression_compute_mse()	26
2.5.2.11 LassoRegression_default_loss()	27
2.5.2.12 LassoRegression_exact_optimum()	27
2.5.2.13 LinearRegression()	28
2.5.2.14 LinearRegression_compute_gradient()	28
2.5.2.15 LinearRegression_compute_mse()	29
2.5.2.16 LinearRegression_default_loss()	29
2.5.2.17 LinearRegression_exact_optimum()	29
2.5.2.18 LinearRegression_free()	30
2.5.2.19 LinearRegression_predict()	30
2.5.2.20 LinearRegression_set_loss()	30
2.5.2.21 LinearRegression_set_mode()	31
2.5.2.22 LinearRegression_train()	31
2.5.2.23 LogisticRegression()	31
2.5.2.24 LogisticRegression_predict()	32
2.5.2.25 LogisticRegression_train()	32

2.5.2.26 RidgeRegression()	33
2.5.2.27 RidgeRegression_compute_gradient()	33
2.5.2.28 RidgeRegression_compute_mse()	34
2.5.2.29 RidgeRegression_default_loss()	34
2.5.2.30 RidgeRegression_exact_optimum()	34
2.5.2.31 Supervised_find_nearest_centroid()	35
2.5.2.32 Supervised_read_csv()	35
2.6 src/unsupervised.c File Reference	36
2.6.1 Macro Definition Documentation	36
2.6.1.1 EPOCHS	36
2.6.1.2 ERROR	37
2.6.1.3 LEARNING_RATE	37
2.6.1.4 MINIBATCH_SIZE	37
2.6.1.5 TOLERANCE	37
2.6.2 Function Documentation	37
2.6.2.1 KMeans()	37
2.6.2.2 KMeans_free()	38
2.6.2.3 KMeans_predict()	38
2.6.2.4 KMeans_train()	38
2.6.2.5 Unsupervised_read_csv()	39
Index	41

# **Chapter 1**

# File Index

## 1.1 File List

Here is a list of all files with brief descriptions:

src/dimreduction.c																						3
src/ <mark>main.c</mark>	 									 												4
src/matrix.c	 									 												4
src/neuralnetwork.c										 												19
src/supervised.c .	 									 												20
src/unsupervised.c										 				 								36

2 File Index

## **Chapter 2**

## **File Documentation**

## 2.1 src/dimreduction.c File Reference

```
#include "dimreduction.h"
```

#### **Functions**

- Matrix PCA\_Reduce (const Matrix \*X, unsigned int target\_rank)

  Computes a low-rank approximation of a matrix using Principle Component Analysis.
- Matrix ChiSquared\_Reduce (const Matrix \*X, const Matrix \*y, unsigned int target\_features)

  Selects the most relevant input features using the chi-squared statistical test.

## 2.1.1 Function Documentation

## 2.1.1.1 ChiSquared\_Reduce()

Selects the most relevant input features using the chi-squared statistical test.

X	original input data
У	output labels
target_features	Number of features to select

#### Returns

Matrix with target\_features input features

## 2.1.1.2 PCA\_Reduce()

```
Matrix PCA_Reduce ( {\tt const~Matrix~*~\it X,} {\tt unsigned~int~\it target\_rank~})
```

Computes a low-rank approximation of a matrix using Principle Component Analysis.

#### **Parameters**

X	the Matrix
target_rank	rank of the output Matrix

#### Returns

Matrix of rank target\_rank

## 2.2 src/main.c File Reference

```
#include <stdio.h>
#include "matrix.h"
```

## **Functions**

• int main ()

## 2.2.1 Function Documentation

## 2.2.1.1 main()

```
int main ( )
```

## 2.3 src/matrix.c File Reference

```
#include "matrix.h"
#include <stdio.h>
#include <stdlib.h>
#include <math.h>
```

#### **Macros**

• #define ERROR(fmt, ...)

Macro to output message to error stream.

#### **Functions**

void Matrix\_free (Matrix mat)

Frees the memory allocated for the Matrix.

void Matrix\_reset (Matrix \*mat)

Sets all entries of the Matrix to 0.

void Matrix\_display (const Matrix \*mat)

Prints the contents of the mAtrix to stdout, formatted with newlines.

bool Matrix\_equal (const Matrix \*mat1, const Matrix \*mat2)

Checks if two Matrices are pairwise (exactly) equal.

bool Matrix approx equal (const Matrix \*mat1, const Matrix \*mat2, double tolerance)

Checks if two Matrices are pairwise equal within a given tolerance.

double Matrix det (const Matrix \*mat)

Computes the determinant of a square Matrix.

double Matrix trace (const Matrix \*mat)

Computes the sum of the diagonal entries of a square Matrix.

double Matrix norm (const Matrix \*mat)

Computes the I2-norm of a Matrix.

double Matrix\_frobenius\_norm (const Matrix \*mat)

Computes the frobenius norm of a Matrix.

double Vector\_norm (const Matrix \*mat, unsigned int k)

Computes the I-k norm of a Vector.

• double Vector\_max (const Matrix \*mat)

Computes the largest value in a row vector.

int Vector\_max\_index (const Matrix \*mat)

Computes the index of the largest value in a row vector.

• Matrix Matrix\_zeros (int rows, int cols)

Initialises new Matrix with entries set to 0.

Matrix Matrix\_identity (int size)

Initialises an Identity matrix (square matrix with diagonal entries 1)

• Matrix Matrix\_from\_array (int rows, int cols, double \*data)

Initialises a Matrix given a 2D array of elements.

Matrix Matrix\_scale (double c, const Matrix \*mat)

Scales a matrix by a constant factor.

Matrix Matrix\_add (const Matrix \*mat1, const Matrix \*mat2)

Sums the entries in two Matrices element-wise.

Matrix Matrix\_sub (const Matrix \*mat1, const Matrix \*mat2)

Subtracts the entries in two Matrices element-wise.

Matrix Matrix multiply (const Matrix \*mat1, const Matrix \*mat2)

Multiplies two Matrices.

Matrix Matrix\_minor (const Matrix \*mat, int row, int col)

Computes the minor of a Matrix by removing a row and a column.

Matrix Matrix row (const Matrix \*mat, int row index)

Get a single row from a matrix.

Matrix Matrix\_col (const Matrix \*mat, int col\_index)

Get a single column from a matrix.

• Matrix Matrix\_slice\_rows (const Matrix \*mat, int start, int end)

Get rows within the specified range.

• Matrix Matrix\_submatrix (const Matrix \*mat, int start\_row, int end\_row, int start\_col, int end\_col)

Get a Matrix that contains a subset of the rows and columns.

• Matrix Matrix\_transpose (const Matrix \*mat)

Flip the rows and columns of a Matrix.

• Matrix Matrix clone (const Matrix \*mat)

Create a new matrix with identical entries.

Matrix Matrix\_inverse (const Matrix \*mat)

Compute the inverse of a square Matrix.

Matrix Matrix\_solve (const Matrix \*A, const Matrix \*b)

Return the solution to the equation Ax=b.

• SVDResult Matrix\_svd (const Matrix \*mat)

Compute the Singular Value Decomposition (SVD) of a Matrix.

#### **Variables**

• const int MATRIX MAX ITER = 1000

Maximum number of iterations for any numerical method on Matrices.

• const double MATRIX\_TOLERANCE = 1e-6

Maximum margin of error for floating point arithmetic on Matrices.

## 2.3.1 Macro Definition Documentation

## 2.3.1.1 ERROR

Macro to output message to error stream.

## 2.3.2 Function Documentation

#### 2.3.2.1 Matrix add()

Sums the entries in two Matrices element-wise.

#### **Parameters**

mat1	first Matrix
mat2	second Matrix

#### Returns

Result of the Matrix addition

## **Exceptions**

Exception	when the Matrices have different dimensions
-----------	---

## 2.3.2.2 Matrix\_approx\_equal()

Checks if two Matrices are pairwise equal within a given tolerance.

#### **Parameters**

mat1	first Matrix
mat2	second Matrix
tolerance	Maximum difference between each element

## Returns

true if the matrices have the same dimension, and the entries are within the tolerance of each other false otherwise

## 2.3.2.3 Matrix\_clone()

Create a new matrix with identical entries.

mat	originial Matrix

#### Returns

Identical Matrix

## 2.3.2.4 Matrix\_col()

Get a single column from a matrix.

#### **Parameters**

mat	original Matrix
col	Index of column to fetch

## Returns

Column vector (in Matrix form) containing the desired column

## **Exceptions**

Exception requires 0 <= col < Matrix.cols
---

## 2.3.2.5 Matrix\_det()

Computes the determinant of a square Matrix.

## **Parameters**

mat	the Matrix

## Returns

double Value of the determinant

## **Exceptions**

Exception for non-square Matrices
-----------------------------------

## 2.3.2.6 Matrix\_display()

Prints the contents of the mAtrix to stdout, formatted with newlines.

#### **Parameters**

```
mat Matrix to be displayed
```

## 2.3.2.7 Matrix\_equal()

Checks if two Matrices are pairwise (exactly) equal.

#### **Parameters**

mat1	first Matrix
mat2	second Matrix

## Returns

true if the matrices have the same dimension, and have equal entries false otherwise

## 2.3.2.8 Matrix\_free()

Frees the memory allocated for the Matrix.

#### **Parameters**

mat | Matrix to be freed

Note

Subsequent accessing of Matrix data causes segfault

## 2.3.2.9 Matrix\_frobenius\_norm()

Computes the frobenius norm of a Matrix.

#### **Parameters**

<i>mat</i> th	e Matrix
---------------	----------

#### Returns

double value of the frobenius norm

## 2.3.2.10 Matrix\_from\_array()

Initialises a Matrix given a 2D array of elements.

#### **Parameters**

rows	Number of rows in the Matrix
cols	Number of columns in the Matrix
data	2D Array, containing the rows of the Matrix

#### Returns

Matrix

## 2.3.2.11 Matrix\_identity()

```
Matrix Matrix_identity ( int \ size \ )
```

Initialises an Identity matrix (square matrix with diagonal entries 1)

## **Parameters**

size	Number of rows and columns
------	----------------------------

## Returns

Matrix

## 2.3.2.12 Matrix\_inverse()

Compute the inverse of a square Matrix.

#### **Parameters**

∣ <i>mat</i> ∣ original Matrix	mat	original Matrix
--------------------------------	-----	-----------------

## Returns

Matrix

## **Exceptions**

```
Exception for non-square Matrices, and for singular Matrices
```

## 2.3.2.13 Matrix\_minor()

Computes the minor of a Matrix by removing a row and a column.

mat	original Matrix
row	Index of row to remove
col	Index of column to remove

#### Returns

Remaining Matrix minor

## **Exceptions**

```
Exception requires 0 <= row < Matrix.rows && 0 <= col < Matrix.cols
```

## 2.3.2.14 Matrix\_multiply()

Multiplies two Matrices.

## **Parameters**

mat1	first Matrix
mat2	second Matrix

## Returns

Result of Matrix multiplication

## **Exceptions**

```
Exception requires mat1->cols == mat2->rows
```

## 2.3.2.15 Matrix\_norm()

Computes the I2-norm of a Matrix.

#### **Parameters**

mat	the Matrix

## Returns

double value of the I2-norm

## 2.3.2.16 Matrix\_reset()

Sets all entries of the Matrix to 0.

#### **Parameters**

```
mat Matrix to be reset
```

## 2.3.2.17 Matrix\_row()

Get a single row from a matrix.

#### **Parameters**

mat	original Matrix	
row	Index of row to fetch	

#### Returns

Row vector (in Matrix form) containing the desired row

## **Exceptions**

```
Exception requires 0 <= row < Matrix.rows
```

## 2.3.2.18 Matrix\_scale()

```
Matrix Matrix_scale ( \label{eq:const_double} \begin{tabular}{ll} $\operatorname{double}\ c$, \\ $\operatorname{const}\ \operatorname{Matrix}\ *\ \operatorname{\it mat}\ ) \end{tabular}
```

Scales a matrix by a constant factor.

С	Factor
mat	original Matrix

#### Returns

New Matrix

## 2.3.2.19 Matrix\_slice\_rows()

Get rows within the specified range.

## **Parameters**

mat	Original matrix
start	Index to start fetching from (inclusive)
end	Index to stop fetching (exclusive)

## Returns

Sub-matrix containing specified rows

## **Exceptions**

```
Exception requires 0 <= start < end <= matrix.rows
```

## 2.3.2.20 Matrix\_solve()

Return the solution to the equation Ax=b.

#### **Parameters**

Α	Matrix A
b	column Vector b (embedded as a Matrix)

## Returns

Column vector x as a Matrix, the solution to the equation

## **Exceptions**

Exception for incompatible dimensions, or singular M	latrices A
--	------------

## 2.3.2.21 Matrix\_sub()

Subtracts the entries in two Matrices element-wise.

#### **Parameters**

mat1	first Matrix
mat2	second Matrix

#### Returns

Result of the Matrix subtraction

## **Exceptions**

eption when the Matrices have diff	ferent dimensions
------------------------------------	-------------------

## 2.3.2.22 Matrix\_submatrix()

Get a Matrix that contains a subset of the rows and columns.

mat	original Matrix
start_row	Index to start fetching rows from (inclusive)
end_row	Index to stop fetching rows from (exclusive)
start_col	Index to start fetching cols from (inclusive)
end_col	Index to stop fetching cols from (exclusive)

#### Returns

Matrix containing the specified rows and columns

## **Exceptions**

```
Exception requires 0 <= start_row < end_row <= mat->rows && 0 <= start_col < end_col <= mat->cols
```

#### 2.3.2.23 Matrix\_svd()

Compute the Singular Value Decomposition (SVD) of a Matrix.

#### **Parameters**

```
mat Matrix
```

## Returns

SVDResult containing orthonormal Matrices U, V, and diagonal Matrix Sigma with entries in decreasing order

## 2.3.2.24 Matrix\_trace()

Computes the sum of the diagonal entries of a square Matrix.

## **Parameters**

```
mat the Matrix
```

#### Returns

double Value of the trace

## **Exceptions**

Exception f	or non-square Matrices
-------------	------------------------

## 2.3.2.25 Matrix\_transpose()

Flip the rows and columns of a Matrix.

## **Parameters**

```
mat original matrix
```

## Returns

Transposed Matrix

## 2.3.2.26 Matrix\_zeros()

Initialises new Matrix with entries set to 0.

#### **Parameters**

rows	Number of rows in the Matrix
cols	Number of columns in the Matrix

## Returns

Matrix

## 2.3.2.27 Vector\_max()

Computes the largest value in a row vector.

mat	the Vector (embedded in a Matrix)

#### Returns

double Value of the largest element in the Vector

## **Exceptions**

Exception	for Matrices that are not row vectors (Must have dimension (Nx1))
-----------	---

## 2.3.2.28 Vector\_max\_index()

Computes the index of the largest value in a row vector.

## Parameters

#### **Returns**

int Index of the largest element in the Vector

## Exceptions

## 2.3.2.29 Vector\_norm()

```
double Vector_norm (  {\rm const\ Matrix\ *\ mat,}  unsigned int k )
```

Computes the I-k norm of a Vector.

## **Parameters**

Ī	mat	the Vector (embedded in a Matrix)
	k	Non-negative integer to adjust norm type

#### **Returns**

double I-k norm

## 2.3.3 Variable Documentation

## 2.3.3.1 MATRIX\_MAX\_ITER

```
const int MATRIX_MAX_ITER = 1000
```

Maximum number of iterations for any numerical method on Matrices.

## 2.3.3.2 MATRIX\_TOLERANCE

```
const double MATRIX_TOLERANCE = 1e-6
```

Maximum margin of error for floating point arithmetic on Matrices.

## 2.4 src/neuralnetwork.c File Reference

```
#include "neuralnetwork.h"
```

## **Functions**

- ActivationFunction NN\_get\_sigmoid ()
- Neuron NN\_create\_neuron (int input\_size, ActivationFunction activation)
- Layer NN\_create\_layer (int num\_neurons, int input\_size, ActivationFunction activation)
- NeuralNetwork NN\_create\_network (int \*layer\_sizes, int num\_layers, ActivationFunction activation)
- Matrix NN\_forward\_pass (NeuralNetwork \*network, const Matrix \*input)
- void NN\_backward\_pass (NeuralNetwork \*network, const Matrix \*input, const Matrix \*expected\_output, double learning\_rate)

## 2.4.1 Function Documentation

#### 2.4.1.1 NN backward pass()

## 2.4.1.2 NN\_create\_layer()

```
Layer NN_create_layer (
                int num_neurons,
                int input_size,
                ActivationFunction activation )
```

## 2.4.1.3 NN\_create\_network()

```
NeuralNetwork NN_create_network (
    int * layer_sizes,
    int num_layers,
    ActivationFunction activation )
```

## 2.4.1.4 NN\_create\_neuron()

```
Neuron NN_create_neuron ( int \ input\_size, ActivationFunction \ activation )
```

## 2.4.1.5 NN\_forward\_pass()

## 2.4.1.6 NN\_get\_sigmoid()

```
ActivationFunction NN_get_sigmoid ( )
```

## 2.5 src/supervised.c File Reference

```
#include "supervised.h"
```

## **Macros**

• #define ERROR(fmt, ...)

Macro to output error message to stdout.

#### **Functions**

int Supervised\_find\_nearest\_centroid (const double \*point, const Matrix \*centroids)

Find the nearest centroid to a given point (embedded as a double array)

LabelledData Supervised\_read\_csv (const char \*filename)

Load Comma Separated Values (CSV) into a LabelledData structure, assuming the last element is the label.

• LinearRegressionModel LinearRegression (const Matrix \*X, const Matrix \*y)

Construct a Linear Regression model with no regularisation.

LinearRegressionModel RidgeRegression (const Matrix \*X, const Matrix \*y, double lambda)

Construct a Linear Regression model with Ridge Regularisation.

LinearRegressionModel LassoRegression (const Matrix \*X, const Matrix \*y, double lambda)

Construct a Linear Regression model with Lasso Regularisation.

void LinearRegression\_train (LinearRegressionModel \*model)

Train a linear regression model.

void LinearRegression set mode (LinearRegressionModel \*model, enum ComputationMode mode)

Set the computation mode of a Linear Regression Model.

void LinearRegression free (LinearRegressionModel model)

Free allocated memory for a LinearRegressionModel.

• void LinearRegression\_set\_loss (LinearRegressionModel \*model, LossFunction loss function)

Change the loss function.

LossFunction LinearRegression\_default\_loss ()

Mean Squared Error (MSE) LossFunction with no regularisation.

LossFunction RidgeRegression default loss ()

Mean Squared Error (MSE) LossFunction with Ridge regularisation (hyper-parameter lambda = model->hyper\_← params[0][0])

LossFunction LassoRegression\_default\_loss ()

Mean Squared Error (MSE) LossFunction with Lasso regularisation (hyper-parameter lambda = model->hyper\_← params[0][0])

• Matrix LinearRegression predict (const LinearRegressionModel \*model, const Matrix \*x new)

Calculate the Linear Regression Estimate for new data.

double LinearRegression\_compute\_mse (const Matrix \*X, const Matrix \*y, const Matrix \*params, const Matrix \*hyper\_params)

Compute Mean Squared Estimate (MSE)

double RidgeRegression\_compute\_mse (const Matrix \*X, const Matrix \*y, const Matrix \*params, const Matrix \*hyper\_params)

Compute Mean Squared Estimate (MSE) with Ridge Regularisation.

double LassoRegression\_compute\_mse (const Matrix \*X, const Matrix \*y, const Matrix \*params, const Matrix \*hyper\_params)

Compute Mean Squared Estimate (MSE) with Lasso Regularisation.

• Matrix LinearRegression\_exact\_optimum (const Matrix \*X, const Matrix \*y, const Matrix \*hyper\_params)

Computes the global optimum algebraically for Linear Regression with no regularisation.

Matrix RidgeRegression\_exact\_optimum (const Matrix \*X, const Matrix \*y, const Matrix \*hyper\_params)

Computes the global optimum algebraically for Linear Regression with Ridge regularisation.

• Matrix LassoRegression exact optimum (const Matrix \*X, const Matrix \*y, const Matrix \*hyper params)

Placeholder function, throws exception since exact optimum can't be computed for Lasso Regression.

 Matrix LinearRegression\_compute\_gradient (const Matrix \*X, const Matrix \*y, const Matrix \*params, const Matrix \*hyper\_params)

Computes the gradient of the MSE Loss Function.

 Matrix RidgeRegression\_compute\_gradient (const Matrix \*X, const Matrix \*y, const Matrix \*params, const Matrix \*hyper\_params)

Compute a sub-gradient of the MSE Loss Function with Ridge regularisation.

 Matrix LassoRegression\_compute\_gradient (const Matrix \*X, const Matrix \*y, const Matrix \*params, const Matrix \*hyper\_params)

Computes a sub-gradient of the MSE Loss Function with Lasso regularisation.

KNNModel KNNClassifier (unsigned int k, const Matrix \*X, const Matrix \*y)

Constructs a KNN model using the majority-vote prediction.

KNNModel KNNRegressor (unsigned int k, const Matrix \*X, const Matrix \*y)

Constructs a KNN model using the majority-vote prediction.

void KNN\_free (KNNModel model)

Free memory allocated by KNN Model.

void KNN\_append\_data (KNNModel \*model, const Matrix \*X\_new, const Matrix \*y\_new)

Add data to existing KNN model.

Matrix KNN\_predict (const KNNModel \*model, const Matrix \*x\_new)

Predict the labels of unseen data using the KNN algorithm.

• LogisticRegressionModel LogisticRegression (const Matrix \*X, const Matrix \*y)

Construct a Logistic Regression Model.

void LogisticRegression\_train (LogisticRegressionModel \*model)

Train a Logistic Regression model.

Matrix LogisticRegression\_predict (const LogisticRegressionModel \*model, const Matrix \*X\_new)

Predict the labels of unseen data.

GaussianNBCModel GaussianNBC (const Matrix \*X, const Matrix \*y)

Create a Naive Bayes Classifier (NBC) model.

• Matrix GaussianNBC\_predict (const GaussianNBCModel \*model, const Matrix \*X\_new)

Use the classifier to predict the labels of unseen data.

#### 2.5.1 Macro Definition Documentation

#### 2.5.1.1 ERROR

Macro to output error message to stdout.

#### 2.5.2 Function Documentation

#### 2.5.2.1 GaussianNBC()

Create a Naive Bayes Classifier (NBC) model.

## **Parameters**

X	Input features
У	Input labels

#### Returns

GaussianNBCModel

## 2.5.2.2 GaussianNBC\_predict()

Use the classifier to predict the labels of unseen data.

#### **Parameters**

model	Model to be used
X_new	Unseen data

## Returns

Matrix Predicted labels in Vector form

## 2.5.2.3 KNN\_append\_data()

Add data to existing KNN model.

model	Existing model
X	New input features
У	New input labels

## 2.5.2.4 KNN\_free()

```
void KNN_free (
          KNNModel model )
```

Free memory allocated by KNN Model.

## **Parameters**

```
model
```

## 2.5.2.5 KNN\_predict()

```
Matrix KNN_predict (  {\rm const~KNNModel~*~model,}   {\rm const~Matrix~*~x\_new~)}
```

Predict the labels of unseen data using the KNN algorithm.

## **Parameters**

model	
x_new	

#### Returns

Matrix

## 2.5.2.6 KNNClassifier()

Constructs a KNN model using the majority-vote prediction.

k	Number of neighbours used by the model
X	Input features
У	Input labels

#### Returns

KNNModel

## 2.5.2.7 KNNRegressor()

Constructs a KNN model using the majority-vote prediction.

## **Parameters**

k Number of neighbours used by the mo	
X	Input features
У	Input labels

## Returns

KNNModel

## 2.5.2.8 LassoRegression()

Construct a Linear Regression model with Lasso Regularisation.

## **Parameters**

X	input data in a Matrix
y output labels in a Vector (embedded as a Matrix	
lambda	Hyper-parameter for Ridge loss

## Returns

LinearRegressionModel with default computation mode set to BATCH

## **Exceptions**

Exception	when X->rows != y->rows
-----------	-------------------------

#### Note

No algebraic solution exists for lasso regression Input Matrix is padded with a 1-column for the bias term

## 2.5.2.9 LassoRegression\_compute\_gradient()

Computes a sub-gradient of the MSE Loss Function with Lasso regularisation.

#### **Parameters**

X	Input features
У	Input labels
params	Current parameters
hyper_params	Contains lambda in data[0][0]

## Returns

Matrix derivative of the loss function wrt the parameters

## 2.5.2.10 LassoRegression\_compute\_mse()

Compute Mean Squared Estimate (MSE) with Lasso Regularisation.

X	Matrix with input data	
У	Vector (embedded in a Matrix) with correct labels	
params	Weights and bias used to make prediction	
hyper_params	Contains lambda in data[0][0]	

#### Returns

double Mean Squared Error

## 2.5.2.11 LassoRegression\_default\_loss()

```
LossFunction LassoRegression_default_loss ( )
```

Mean Squared Error (MSE) LossFunction with Lasso regularisation (hyper-parameter lambda = model->hyper $_{\leftarrow}$  params[0][0])

#### Returns

LossFunction

Note

Cannot be trained with computation mode set to ALGEBRAIC

## 2.5.2.12 LassoRegression\_exact\_optimum()

```
Matrix LassoRegression_exact_optimum (  {\rm const\ Matrix\ *\ } X,   {\rm const\ Matrix\ *\ } y,   {\rm const\ Matrix\ *\ } hyper\_params\ )
```

Placeholder function, throws exception since exact optimum can't be computed for Lasso Regression.

#### **Parameters**

X	Input features
У	Input labels
hyper_params	Contains lambda in data[0][0]

#### Returns

Matrix Global optimum value of params

## **Exceptions**

Exception

## 2.5.2.13 LinearRegression()

Construct a Linear Regression model with no regularisation.

## **Parameters**

Χ	input data in a Matrix	
У	output labels in a Vector (embedded as a Matrix)	

#### Returns

LinearRegressionModel with default computation mode set to ALGEBRAIC

## **Exceptions**

```
Exception when X->rows != y->rows
```

#### Note

Input Matrix is padded with a 1-column for the bias term

## 2.5.2.14 LinearRegression\_compute\_gradient()

Computes the gradient of the MSE Loss Function.

## **Parameters**

X	Input features
У	Input labels
params	Current parameters
hyper_params	Unneeded

#### Returns

Matrix derivative of the loss function wrt the parameters

## 2.5.2.15 LinearRegression\_compute\_mse()

Compute Mean Squared Estimate (MSE)

#### **Parameters**

X	Matrix with input data	
У	Vector (embedded in a Matrix) with correct labels	
params	Weights and bias used to make prediction	
hyper_params	Unneeded	

#### Returns

double Mean Squared Error

## 2.5.2.16 LinearRegression\_default\_loss()

```
LossFunction LinearRegression_default_loss ( )
```

Mean Squared Error (MSE) LossFunction with no regularisation.

## Returns

LossFunction

## 2.5.2.17 LinearRegression\_exact\_optimum()

Computes the global optimum algebraically for Linear Regression with no regularisation.

X	Input features
У	Input labels
hyper params	Unneeded

#### Returns

Matrix Global optimum value of params

## 2.5.2.18 LinearRegression\_free()

```
\begin{tabular}{ll} {\tt void LinearRegression\_free (} \\ {\tt LinearRegressionModel } {\it model )} \end{tabular}
```

Free allocated memory for a LinearRegressionModel.

#### **Parameters**

model Model to be freed
-------------------------

#### Note

Further calls to model may segfault

## 2.5.2.19 LinearRegression\_predict()

Calculate the Linear Regression Estimate for new data.

#### **Parameters**

model	Linear Regression Model to predict with	
x_new	Matrix containing the new (unpadded) data as rows	

#### Returns

Vector (embedded as Matrix) containing the predicted labels

## 2.5.2.20 LinearRegression set loss()

Change the loss function.

#### **Parameters**

model	model to be changed
loss_function	New loss function

#### Note

model needs to be retrained before predictions are made

## 2.5.2.21 LinearRegression\_set\_mode()

Set the computation mode of a Linear Regression Model.

#### **Parameters**

model	The model to be changed
mode	Desired computation mode

## Note

Some LossFunctions do not have an algebraic solution, and may throw exceptions when trained with computation mode ALGEBRAIC

## 2.5.2.22 LinearRegression\_train()

Train a linear regression model.

## **Parameters**

```
model Model to be trained
```

# 2.5.2.23 LogisticRegression()

 ${\tt LogisticRegressionModel\ LogisticRegression\ (}$ 

```
const Matrix * X, const Matrix * y)
```

Construct a Logistic Regression Model.

#### **Parameters**

Χ	Input features
У	Input labels

## Returns

LogisticRegressionModel

## 2.5.2.24 LogisticRegression\_predict()

```
Matrix LogisticRegression_predict (  {\tt const\ LogisticRegressionModel\ *\ model,}   {\tt const\ Matrix\ *\ X\_new\ )}
```

Predict the labels of unseen data.

#### **Parameters**

model	Model to be used
X_new	Data to be predicted

## Returns

Vector (embedded in a Matrix) containing the predicted labels

# Exceptions

Exception requires model to be trained
--

# 2.5.2.25 LogisticRegression\_train()

```
void LogisticRegression_train ( {\tt LogisticRegressionModel*\ model})
```

Train a Logistic Regression model.

## **Parameters**

model	to be trained
model	to be trained

## 2.5.2.26 RidgeRegression()

Construct a Linear Regression model with Ridge Regularisation.

#### **Parameters**

X	input data in a Matrix
У	output labels in a Vector (embedded as a Matrix)
lambda	Hyper-parameter for Ridge loss

#### Returns

LinearRegressionModel with default computation mode set to ALGEBRAIC

# **Exceptions**

Exc	ception	when X->rows	!= y->rows
-----	---------	--------------	------------

#### Note

Input Matrix is padded with a 1-column for the bias term

## 2.5.2.27 RidgeRegression\_compute\_gradient()

Compute a sub-gradient of the MSE Loss Function with Ridge regularisation.

## **Parameters**

X	Input features
У	Input labels
params	Current parameters
hyper_params	Contains lambda in data[0][0]

#### Returns

Matrix derivative of the loss function wrt the parameters

# 2.5.2.28 RidgeRegression\_compute\_mse()

Compute Mean Squared Estimate (MSE) with Ridge Regularisation.

#### **Parameters**

X	Matrix with input data
У	Vector (embedded in a Matrix) with correct labels
params	Weights and bias used to make prediction
hyper_params	Contains lambda in data[0][0]

#### Returns

double Mean Squared Error

## 2.5.2.29 RidgeRegression\_default\_loss()

```
LossFunction RidgeRegression_default_loss ( )
```

 $\label{lossFunction} \begin{tabular}{ll} Mean Squared Error (MSE) LossFunction with Ridge regularisation (hyper-parameter lambda = model->hyper\_ constraints and params [0][0]) \\ \end{tabular}$ 

## Returns

LossFunction

#### 2.5.2.30 RidgeRegression\_exact\_optimum()

```
Matrix RidgeRegression_exact_optimum (  {\rm const\ Matrix\ *\ } X,   {\rm const\ Matrix\ *\ } y,   {\rm const\ Matrix\ *\ } hyper\_params\ )
```

Computes the global optimum algebraically for Linear Regression with Ridge regularisation.

#### **Parameters**

X	Input features
У	Input labels
hyper_params	Contains lambda in data[0][0]

#### Returns

Matrix Global optimum value of params

# 2.5.2.31 Supervised\_find\_nearest\_centroid()

Find the nearest centroid to a given point (embedded as a double array)

#### **Parameters**

point	Array of values representing point in euclidean space
centroids	matrix containing centroid coordinates as rows

## Returns

index of closes centroid

## 2.5.2.32 Supervised\_read\_csv()

Load Comma Separated Values (CSV) into a LabelledData structure, assuming the last element is the label.

# **Parameters**

filename

## Returns

LabelledData parsed data

## **Exceptions**

Exception | file must exist, and be structured appropriately

# 2.6 src/unsupervised.c File Reference

#include "unsupervised.h"

#### **Macros**

- #define LEARNING RATE 0.01
- #define EPOCHS 1000
- #define MINIBATCH\_SIZE 32
- #define TOLERANCE 1e-6
- #define ERROR(fmt, ...)

#### **Functions**

• Matrix Unsupervised\_read\_csv (const char \*filename)

Reads a Matrix from a Comma Separated Value (CSV) file.

• KMeansModel KMeans (unsigned int k, const Matrix \*X)

Construct a KMeansModel with a set of data.

• void KMeans\_train (KMeansModel \*model)

Train the model using lloyd's algorithm.

• Matrix KMeans\_predict (const KMeansModel \*model, const Matrix \*X\_new)

Predict the cluster of unseen data using a KMeansModel.

• void KMeans\_free (KMeansModel \*model)

Free memory allocated to KMeans Model.

#### 2.6.1 Macro Definition Documentation

## 2.6.1.1 EPOCHS

#define EPOCHS 1000

# 2.6.1.2 ERROR

# 2.6.1.3 LEARNING\_RATE

```
#define LEARNING_RATE 0.01
```

# 2.6.1.4 MINIBATCH\_SIZE

```
#define MINIBATCH_SIZE 32
```

## **2.6.1.5 TOLERANCE**

```
#define TOLERANCE 1e-6
```

## 2.6.2 Function Documentation

## 2.6.2.1 KMeans()

Construct a KMeansModel with a set of data.

#### **Parameters**

k	Number of clusters
Χ	Data

#### Returns

KMeansModel

# 2.6.2.2 KMeans\_free()

Free memory allocated to KMeans Model.

#### **Parameters**

model Model to be freed
-------------------------

Note

Subsequent accesses to the model may segfault

# 2.6.2.3 KMeans\_predict()

Predict the cluster of unseen data using a KMeansModel.

# **Parameters**

model	
X_new	

## Returns

Matrix Predicted clusters

# 2.6.2.4 KMeans\_train()

Train the model using lloyd's algorithm.

# **Parameters**

model Model to be trained

# 2.6.2.5 Unsupervised\_read\_csv()

Reads a Matrix from a Comma Separated Value (CSV) file.

**Parameters** 

filename

# Returns

Matrix read data

# Index

ChiSquared_Reduce	supervised.c, 27
dimreduction.c, 3	LinearRegression_compute_gradient
	supervised.c, 28
dimreduction.c	LinearRegression_compute_mse
ChiSquared_Reduce, 3	supervised.c, 28
PCA_Reduce, 4	LinearRegression_default_loss
	supervised.c, 29
EPOCHS	LinearRegression_exact_optimum
unsupervised.c, 36	supervised.c, 29
ERROR	LinearRegression free
matrix.c, 6	supervised.c, 30
supervised.c, 22	LinearRegression_predict
unsupervised.c, 36	supervised.c, 30
	LinearRegression_set_loss
GaussianNBC	supervised.c, 30
supervised.c, 22	LinearRegression_set_mode
GaussianNBC_predict	<del>-</del>
supervised.c, 23	supervised.c, 31
	LinearRegression_train
KMeans	supervised.c, 31
unsupervised.c, 37	LogisticRegression
KMeans_free	supervised.c, 31
unsupervised.c, 38	LogisticRegression_predict
KMeans_predict	supervised.c, 32
unsupervised.c, 38	LogisticRegression_train
KMeans_train	supervised.c, 32
unsupervised.c, 38	
KNN_append_data	main . ,
supervised.c, 23	main.c, 4
KNN_free	main.c
supervised.c, 23	main, 4
KNN_predict	matrix.c
supervised.c, 24	ERROR, 6
KNNClassifier	Matrix_add, 6
supervised.c, 24	Matrix_approx_equal, 7
KNNRegressor	Matrix_clone, 7
supervised.c, 25	Matrix_col, 8
,	Matrix_det, 8
LassoRegression	Matrix_display, 9
supervised.c, 25	Matrix_equal, 9
LassoRegression_compute_gradient	Matrix_free, 9
supervised.c, 26	Matrix_frobenius_norm, 10
LassoRegression_compute_mse	Matrix_from_array, 10
supervised.c, 26	Matrix_identity, 10
LassoRegression_default_loss	Matrix_inverse, 11
supervised.c, 27	MATRIX_MAX_ITER, 19
LassoRegression_exact_optimum	Matrix minor, 11
supervised.c, 27	Matrix multiply, 12
LEARNING RATE	Matrix_norm, 12
unsupervised.c, 37	Matrix_reset, 12
LinearRegression	Matrix row, 13
- 3	<u> </u>

42 INDEX

Matrix_scale, 13	matrix.c, 15
Matrix_slice_rows, 14	Matrix_svd
Matrix_solve, 14	matrix.c, 16
Matrix_sub, 15	MATRIX_TOLERANCE
Matrix_submatrix, 15	matrix.c, 19
Matrix_svd, 16	Matrix_trace
MATRIX_TOLERANCE, 19	matrix.c, 16
Matrix_trace, 16	Matrix_transpose
Matrix_transpose, 16	matrix.c, 16
Matrix_zeros, 17	Matrix_zeros
Vector_max, 17	matrix.c, 17
Vector_max_index, 18	MINIBATCH_SIZE
Vector_norm, 18	unsupervised.c, 37
Matrix_add	
matrix.c, 6	neuralnetwork.c
Matrix_approx_equal	NN_backward_pass, 19
matrix.c, 7	NN_create_layer, 19
Matrix_clone	NN_create_network, 20
matrix.c, 7	NN_create_neuron, 20
Matrix_col	NN_forward_pass, 20
matrix.c, 8	NN_get_sigmoid, 20
Matrix_det	NN_backward_pass
matrix.c, 8	neuralnetwork.c, 19
Matrix_display	NN_create_layer
matrix.c, 9	neuralnetwork.c, 19
Matrix_equal	NN_create_network
matrix.c, 9	neuralnetwork.c, 20
Matrix free	NN_create_neuron
matrix.c, 9	neuralnetwork.c, 20
Matrix frobenius norm	NN_forward_pass
matrix.c, 10	neuralnetwork.c, 20
Matrix_from_array	NN_get_sigmoid
matrix.c, 10	neuralnetwork.c, 20
Matrix_identity	
matrix.c, 10	PCA_Reduce
Matrix inverse	dimreduction.c, 4
matrix.c, 11	
MATRIX MAX ITER	RidgeRegression
matrix.c, 19	supervised.c, 33
Matrix minor	RidgeRegression_compute_gradient
matrix.c, 11	supervised.c, 33
Matrix_multiply	RidgeRegression_compute_mse
matrix.c, 12	supervised.c, 34
Matrix norm	RidgeRegression_default_loss
matrix.c, 12	supervised.c, 34
Matrix_reset	RidgeRegression_exact_optimum
matrix.c. 12	supervised.c, 34
•	
Matrix_row	src/dimreduction.c, 3
matrix.c, 13	src/main.c, 4
Matrix_scale	src/matrix.c, 4
matrix.c, 13	src/neuralnetwork.c, 19
Matrix_slice_rows	src/supervised.c, 20
matrix.c, 14	src/unsupervised.c, 36
Matrix_solve	supervised.c
matrix.c, 14	ERROR, 22
Matrix_sub	GaussianNBC, 22
matrix.c, 15	GaussianNBC_predict, 23
Matrix_submatrix	KNN_append_data, 23
	_ · · · _ ·

INDEX 43

```
KNN_free, 23
    KNN predict, 24
    KNNClassifier, 24
    KNNRegressor, 25
    LassoRegression, 25
    LassoRegression compute gradient, 26
    LassoRegression_compute_mse, 26
    LassoRegression default loss, 27
    LassoRegression exact optimum, 27
    LinearRegression, 27
    LinearRegression_compute_gradient, 28
    LinearRegression_compute_mse, 28
    LinearRegression_default_loss, 29
    LinearRegression_exact_optimum, 29
    LinearRegression_free, 30
    LinearRegression_predict, 30
    LinearRegression set loss, 30
    LinearRegression set mode, 31
    LinearRegression_train, 31
    LogisticRegression, 31
    LogisticRegression predict, 32
    LogisticRegression train, 32
     RidgeRegression, 33
     RidgeRegression_compute_gradient, 33
     RidgeRegression compute mse, 34
     RidgeRegression_default_loss, 34
    RidgeRegression_exact_optimum, 34
    Supervised find nearest centroid, 35
    Supervised read csv, 35
Supervised_find_nearest_centroid
    supervised.c, 35
Supervised_read_csv
    supervised.c, 35
TOLERANCE
    unsupervised.c, 37
unsupervised.c
    EPOCHS, 36
    ERROR, 36
    KMeans, 37
    KMeans free, 38
    KMeans predict, 38
    KMeans_train, 38
    LEARNING_RATE, 37
    MINIBATCH SIZE, 37
    TOLERANCE, 37
    Unsupervised_read_csv, 39
Unsupervised_read_csv
    unsupervised.c, 39
Vector_max
    matrix.c, 17
Vector_max_index
    matrix.c, 18
Vector norm
    matrix.c, 18
```