



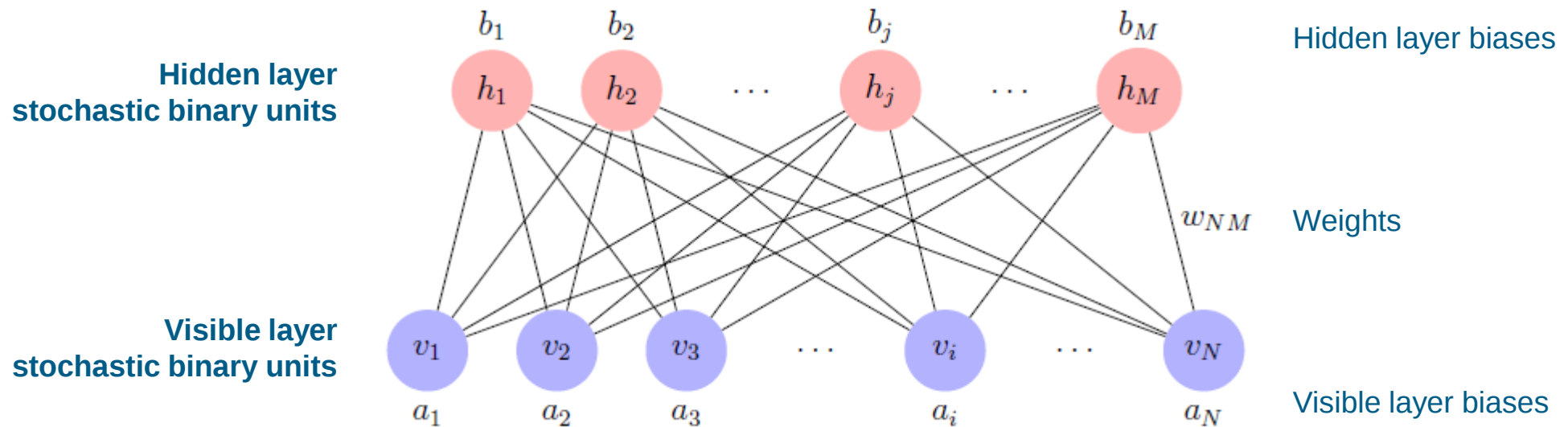
The Market Generator

Alexei Kondratyev

Financial Engineering Workshop, City, University of London, 25 November 2020

Here for good
standardchartered.com

Restricted Boltzmann Machine (RBM)



The probability of configuration (v, h) is given by Boltzmann (Gibbs) distribution:

$$P(v, h) = \frac{1}{Z} e^{-E(v, h)}, \quad Z = \sum_{\{v, h\}} e^{-E(v, h)}$$

where the energy of a configuration (v, h) is given by expression:

$$E(v, h) = \sum_{i=1}^m a_i v_i + \sum_{j=1}^n b_j h_j + \sum_{i=1}^m \sum_{j=1}^n w_{ij} v_i h_j$$

Stochastic activation units

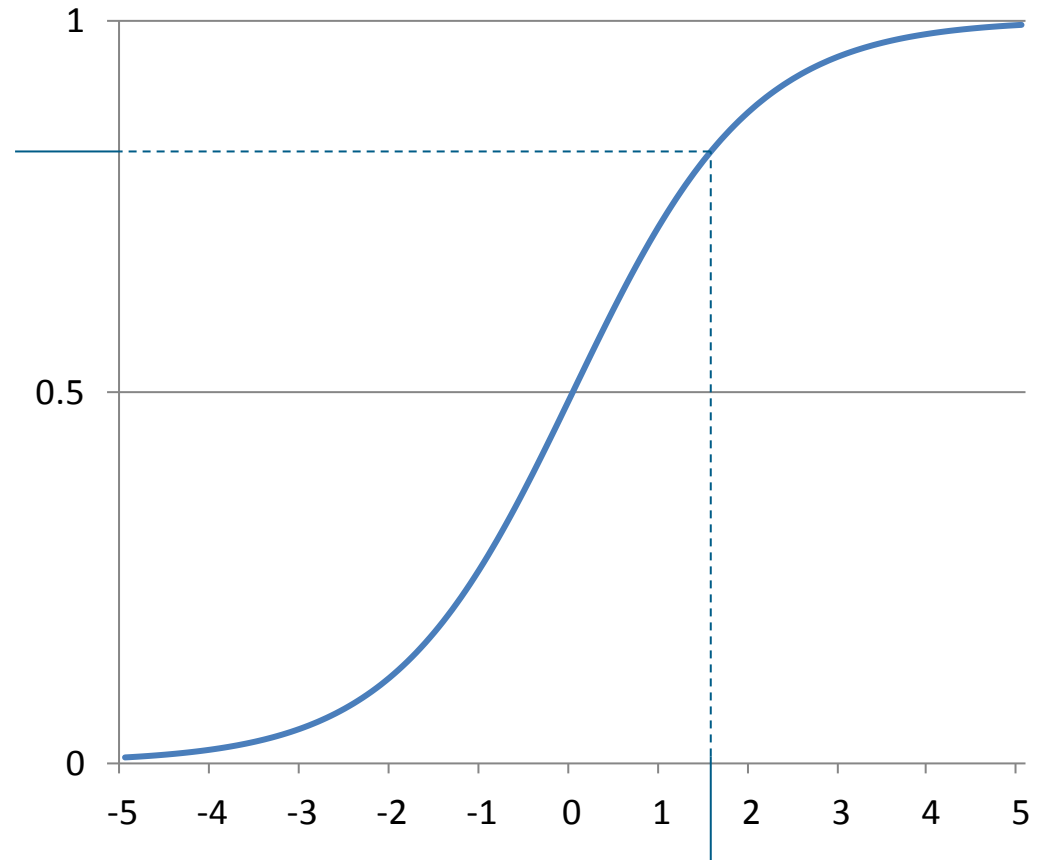
2 Hidden node j
firing probability

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

3 Deciding whether node j
fires or not:

$$z \sim U[0,1]$$

IF $\sigma(x) \geq z$
THEN $h_j = 1$
ELSE $h_j = 0$



1 Aggregated input for
the hidden node j

$$x = b_j + \sum_{i=1}^m w_{ij} v_i$$

Training algorithm for RBM

Marginal probability of visible state:

$$P(v) = \frac{1}{Z} \sum_h e^{-E(v,h)}$$

Weight updates with stochastic gradient ascent
(maximisation of the likelihood of the observed data):

$$\Delta w_{ij} = \eta \frac{\partial \log(P(v))}{\partial w_{ij}}$$

The gradient has the form:

$$\frac{\partial \log(P(v))}{\partial w_{ij}} = \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{model}}$$

where $\langle \dots \rangle$ denotes expectations under the distribution specified by the subscript.

Contrastive Divergence Algorithm

Algorithm 1 k -step contrastive divergence

Result: Weights and biases updates

Input: training minibatch S

Input: model parameters a_i, b_j, w_{ij} , $i = 1, \dots, N$, $j = 1, \dots, M$ (before update)

Initialization: $\forall i, j : \Delta w_{ij} = \Delta a_i = \Delta b_j = 0$

```
for  $\mathbf{v} \in S$  do
   $\mathbf{v}^{(0)} \leftarrow \mathbf{v}$ 
  for  $t = 0, \dots, k - 1$  do
    for  $j = 1, \dots, M$  do
      | sample Bernoulli random variable  $h_j^{(t)} \sim p(h_j | \mathbf{v}^{(t)})$ 
    end
    for  $i = 1, \dots, N$  do
      | sample Bernoulli random variable  $v_i^{(t+1)} \sim p(v_i | \mathbf{h}^{(t)})$ 
    end
  end
  for  $i = 1, \dots, N$ ,  $j = 1, \dots, M$  do
    |  $\Delta w_{ij} \leftarrow \Delta w_{ij} + \eta \left( p(h_j = 1 | \mathbf{v}^{(0)}) v_i^{(0)} - p(h_j = 1 | \mathbf{v}^{(k)}) v_i^{(k)} \right)$ 
  end
  for  $i = 1, \dots, N$  do
    |  $\Delta a_i \leftarrow \Delta a_i + \eta \left( v_i^{(0)} - v_i^{(k)} \right)$ 
  end
  for  $j = 1, \dots, M$  do
    |  $\Delta b_j \leftarrow \Delta b_j + \eta \left( p(h_j = 1 | \mathbf{v}^{(0)}) - p(h_j = 1 | \mathbf{v}^{(k)}) \right)$ 
  end
end
```

Asja Fischer and
Christian Igel (2014)
*Training Restricted Boltzmann
Machines: An Introduction*
Pattern Recognition, Volume 47

Geoffrey Hinton (2010)
*A Practical Guide to Training
Restricted Boltzmann Machines*
Department of Computer Science
University of Toronto
UTML TR 2010-003
[https://www.cs.toronto.edu/~hinton/
absps/guideTR.pdf](https://www.cs.toronto.edu/~hinton/absps/guideTR.pdf)

Transformation of the real-valued datasets

Algorithm 2 Real-valued to integer to binary transformation (training phase)

Result: Conversion of real-valued data set into binary features

Input: $X_{real}^{(n)}(l)$ – a real-valued data set: $l = 1, \dots, N_{samples}$, $n = 1, \dots, N_{variables}$

```
for  $n = 1, \dots, N_{variables}$  do
     $X_{min}^{(n)} \leftarrow \min_{\{l\}}(X_{real}^{(n)}) - \epsilon_{min}^{(n)}$ ,  $\epsilon_{min}^{(n)} \geq 0$ 
     $X_{max}^{(n)} \leftarrow \max_{\{l\}}(X_{real}^{(n)}) + \epsilon_{max}^{(n)}$ ,  $\epsilon_{max}^{(n)} \geq 0$ 
    for  $l = 1, \dots, N_{samples}$  do
         $X_{integer}^{(n)}(l) \leftarrow \text{int} \left( 65535 \times (X_{real}^{(n)}(l) - X_{min}^{(n)}) / (X_{max}^{(n)} - X_{min}^{(n)}) \right)$ 
         $X_{binary}^{(n)}(l) \leftarrow \text{binarize} \left( X_{integer}^{(n)}(l) \right)$ 
    end
end
```

Each data sample is represented by a 16-digit binary number ($2^{16} - 1 = 65535$) with every digit becoming a separate feature. The total number of features is $16 \times N_{variables}$.

Algorithm 3 Binary to integer to real-valued transformation (sampling phase)

Result: Conversion of the generated 16-digit binary sample into real-valued sample

Input: $\hat{X}^{(n)}[m]$ – generated 16-digit binary sample: $m = 0, \dots, 15$, $n = 1, \dots, N_{variables}$

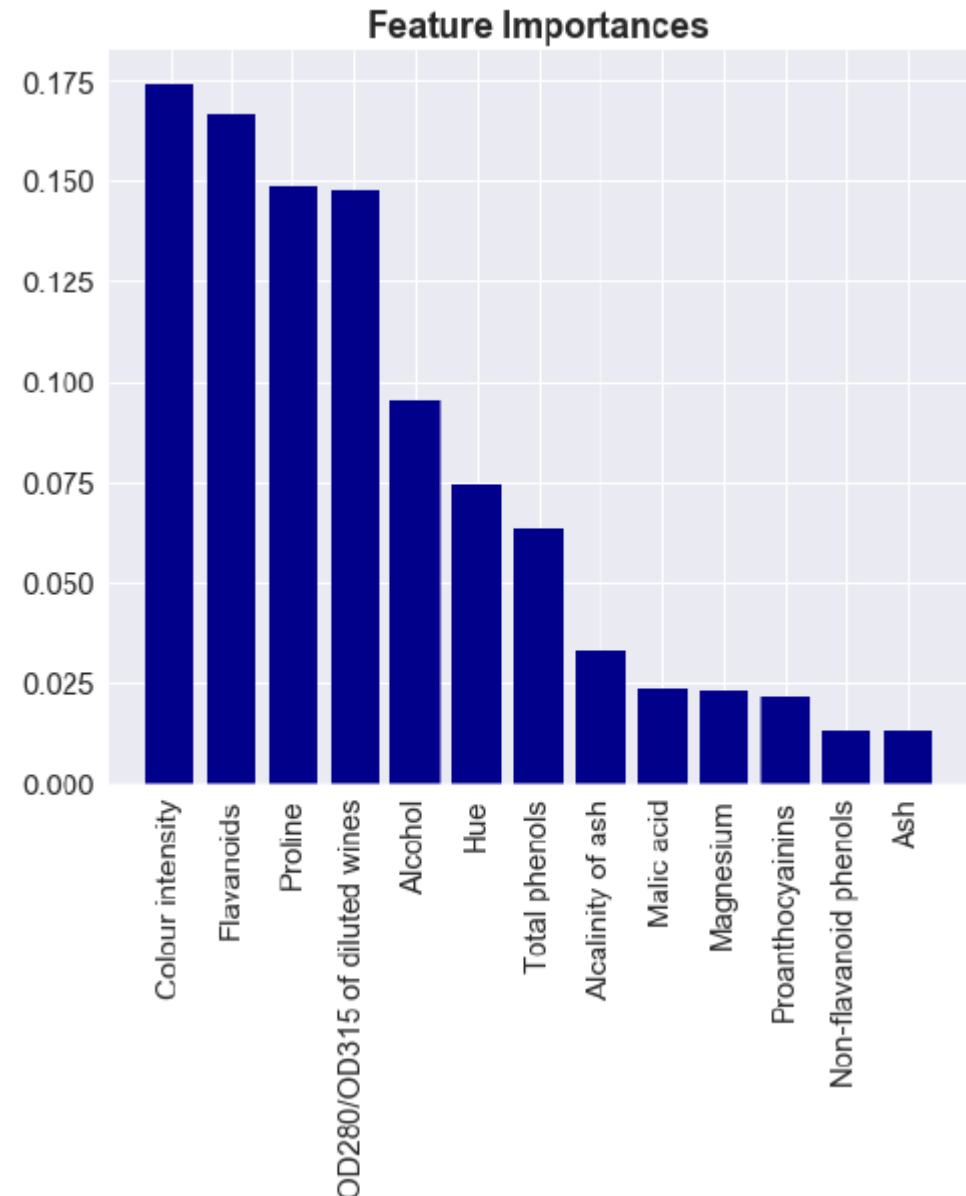
```
for  $n = 1, \dots, N_{variables}$  do
     $\hat{X}_{integer}^{(n)} \leftarrow 0$ 
    for  $m = 0, \dots, 15$  do
         $\hat{X}_{integer}^{(n)} \leftarrow \hat{X}_{integer}^{(n)} + 2^m \times \hat{X}^{(n)}[15 - m]$ 
    end
     $\hat{X}_{real}^{(n)} \leftarrow X_{min}^{(n)} + \hat{X}_{integer}^{(n)} \times (X_{max}^{(n)} - X_{min}^{(n)}) / 65535$ 
end
```

Alexei Kondratyev and
Christian Schwarz (2019)
The Market Generator

<https://ssrn.com/abstract=3384948>

Illustrative example – feature extraction

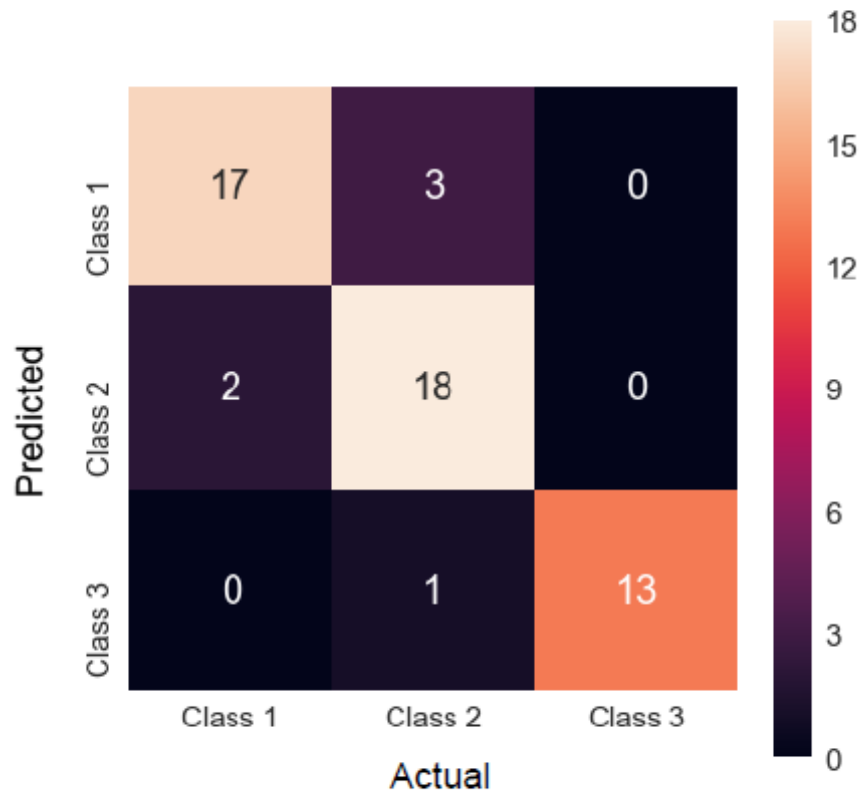
- Open source dataset of wine samples (University of California, Irvine)
<https://archive.ics.uci.edu/ml/machine-learning-databases/wine/>
- 178 wine samples, 13 real-valued features, 3 grape types
- Our plan:
 - 1) Standardise and binarise features
 - 2) Split original dataset into training and validation datasets (70:30)
 - 3) Train MLP Classifier and make predictions for the validation dataset
 - 4) Train combined classifier consisting of RBM and MLP and make predictions for the validation dataset
 - 5) Compare results



Illustrative example – feature extraction

MLP Classifier

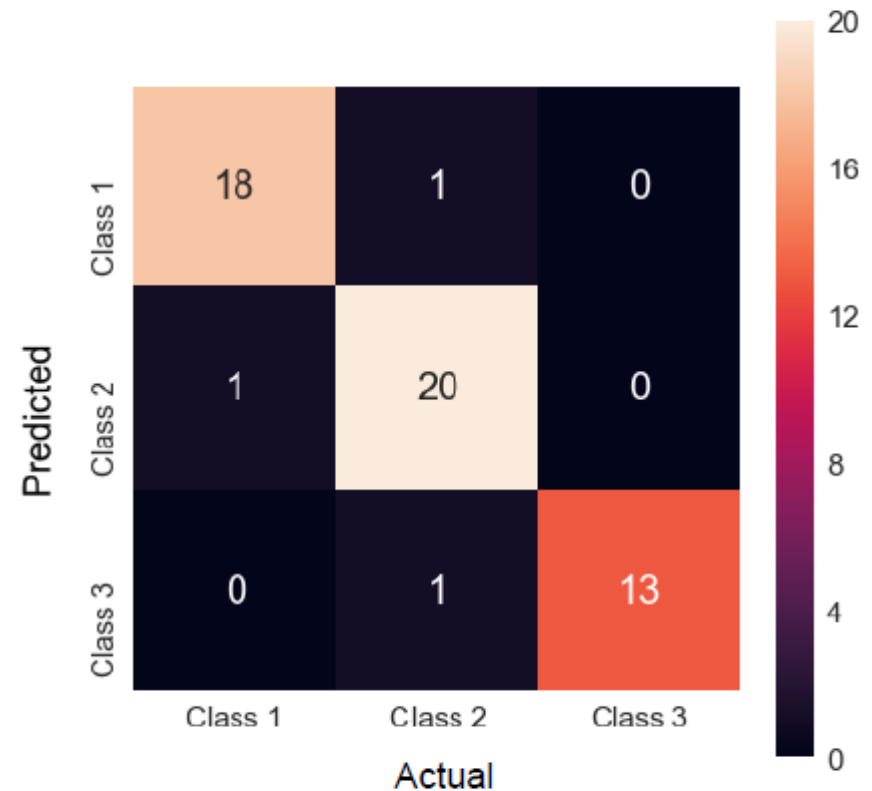
88.9% correct predictions



```
mlpc = MLPClassifier(hidden_layer_sizes=(7, 10), alpha=0.001,\
                    tol=0.0001, max_iter=200, random_state=0)
```

RBM + MLP Classifier

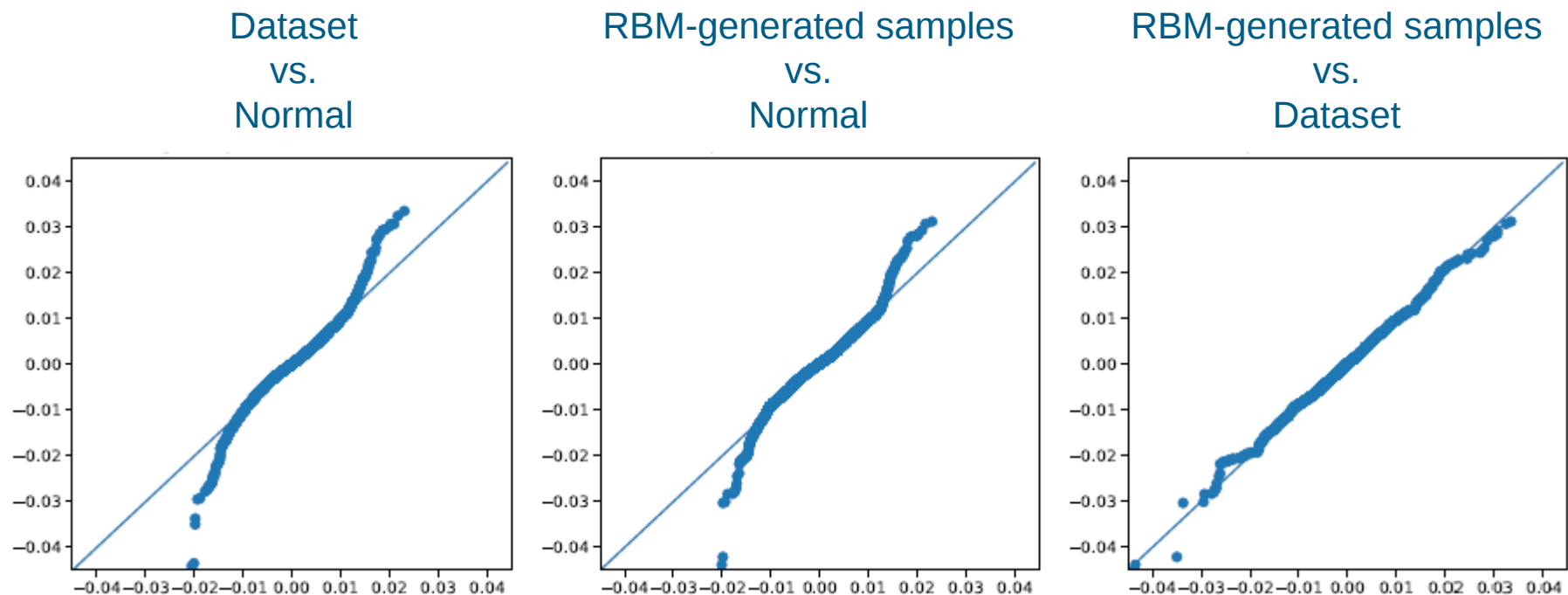
94.4% correct predictions



```
rbm = BernoulliRBM(n_components=7, n_iter=200, random_state=0)
mlpc = MLPClassifier(hidden_layer_sizes=(10), alpha=0.001,\
                    tol=0.0001, max_iter=200, random_state=0)
classifier = Pipeline(steps=[('rbm', rbm), ('mlpc', mlpc)])
```

Generation of synthetic market data

QQ-plots of USDJPY spot FX daily log-returns (dataset: 1999-2019)



Alexei Kondratyev and
Christian Schwarz (2019)
The Market Generator

<https://ssrn.com/abstract=3384948>

Generation of synthetic market data (continued)

Reconstruction of correlations

Currency pairs	data set			RBM-generated samples		
	Pearson	Spearman	Kendall	Pearson	Spearman	Kendall
EURUSD/GBPUSD	63.7%	62.5%	45.3%	55.7% ($\pm 4.4\%$)	59.1% ($\pm 2.2\%$)	42.7% ($\pm 1.8\%$)
EURUSD/USDJPY	-28.2%	-31.2%	-21.8%	-26.0% ($\pm 2.5\%$)	-28.7% ($\pm 1.4\%$)	-20.0% ($\pm 1.0\%$)
EURUSD/USDCAD	-44.3%	-38.6%	-26.8%	-40.1% ($\pm 3.8\%$)	-37.1% ($\pm 1.8\%$)	-25.7% ($\pm 1.3\%$)
GBPUSD/USDJPY	-13.5%	-21.2%	-14.7%	-15.5% ($\pm 5.4\%$)	-21.2% ($\pm 0.9\%$)	-14.6% ($\pm 0.7\%$)
GBPUSD/USDCAD	-43.1%	-35.8%	-24.8%	-38.5% ($\pm 3.2\%$)	-34.3% ($\pm 2.4\%$)	-23.7% ($\pm 1.7\%$)
USDJPY/USDCAD	3.0%	7.4%	5.1%	5.6% ($\pm 2.5\%$)	8.5% ($\pm 1.5\%$)	5.8% ($\pm 1.0\%$)

Reconstruction of tail behaviour

Currency pair	data set		RBM-generated samples	
	1st percentile	99th percentile	1st percentile	99th percentile
EURUSD	-1.58%	1.53%	-1.65% ($\pm 0.05\%$)	1.61% ($\pm 0.05\%$)
GBPUSD	-1.48%	1.34%	-1.55% ($\pm 0.04\%$)	1.53% ($\pm 0.10\%$)
USDJPY	-1.73%	1.66%	-1.84% ($\pm 0.07\%$)	1.78% ($\pm 0.07\%$)
USDCAD	-1.42%	1.50%	-1.47% ($\pm 0.04\%$)	1.55% ($\pm 0.06\%$)

Reconstruction of annualised volatilities

Currency pair	Historical volatility	RBM-generated samples
EURUSD	9.7%	9.8% ($\pm 0.7\%$)
GBPUSD	9.4%	9.5% ($\pm 0.7\%$)
USDJPY	10.2%	10.4% ($\pm 0.4\%$)
USDCAD	8.9%	8.6% ($\pm 0.8\%$)

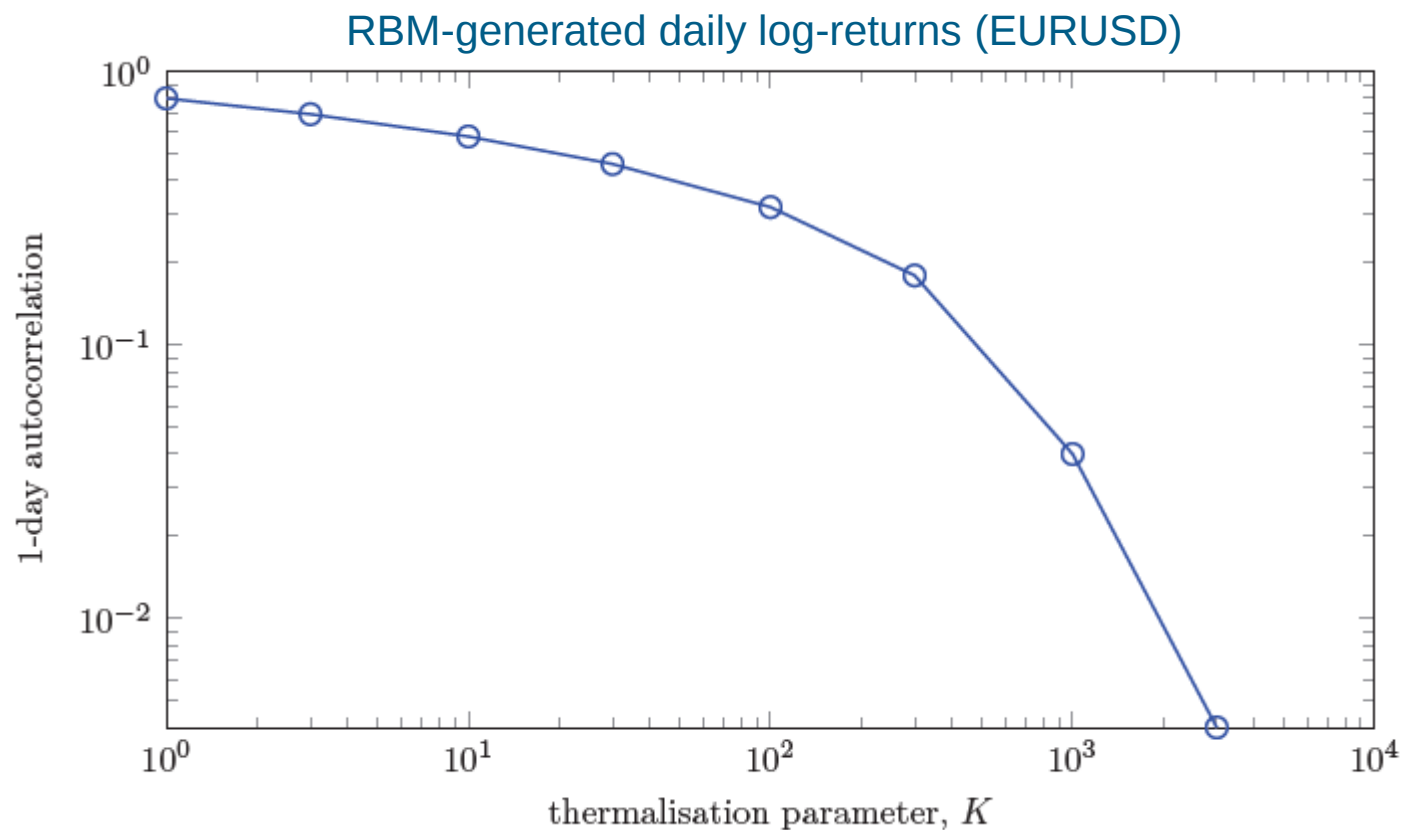
Alexei Kondratyev and
Christian Schwarz (2019)

The Market Generator

<https://ssrn.com/abstract=3384948>

Generation of synthetic market data (continued)

RBM can be used to generate either fully independent or autocorrelated samples with desired degree of autocorrelation.



Alexei Kondratyev and
Christian Schwarz (2019)
The Market Generator

<https://ssrn.com/abstract=3384948>

Generation of synthetic market data (continued)

RBM is able to handle non-stationarity through conditional sampling.

Conditional sampling				
Currency pair	Low volatility environment		High volatility environment	
	Historical volatility	RBM-generated samples	Historical volatility	RBM-generated samples
EURUSD	6.8%	6.7% ($\pm 0.2\%$)	12.2%	10.8% ($\pm 0.5\%$)
GBPUSD	6.9%	6.8% ($\pm 0.3\%$)	13.1%	13.7% ($\pm 0.7\%$)
USDJPY	8.0%	8.9% ($\pm 0.3\%$)	12.8%	12.7% ($\pm 0.6\%$)
USDCAD	6.2%	6.4% ($\pm 0.2\%$)	13.0%	11.6% ($\pm 0.7\%$)

Alexei Kondratyev and
Christian Schwarz (2019)
The Market Generator

<https://ssrn.com/abstract=3384948>

Further Applications

- Applications:
 - Data Anonymisation
 - Outlier Detection
 - Fighting Overfitting
- RBMs can be trained on relatively small datasets, which often are too small to train GANs.
- We apply our model on several standard medical and financial datasets.
- We generate synthetic data (anonymised dataset) with the same statistical properties as the original data (i.e. a potentially sensitive dataset that can be used to deduce patients or clients details).

Alexei Kondratyev, Christian Schwarz and Blanka Horvath (2020)
*Data Anonymisation, Outlier Detection and Fighting Overfitting
with Restricted Boltzmann Machines*
<https://ssrn.com/abstract=3526436>

Example 1: Wisconsin dataset

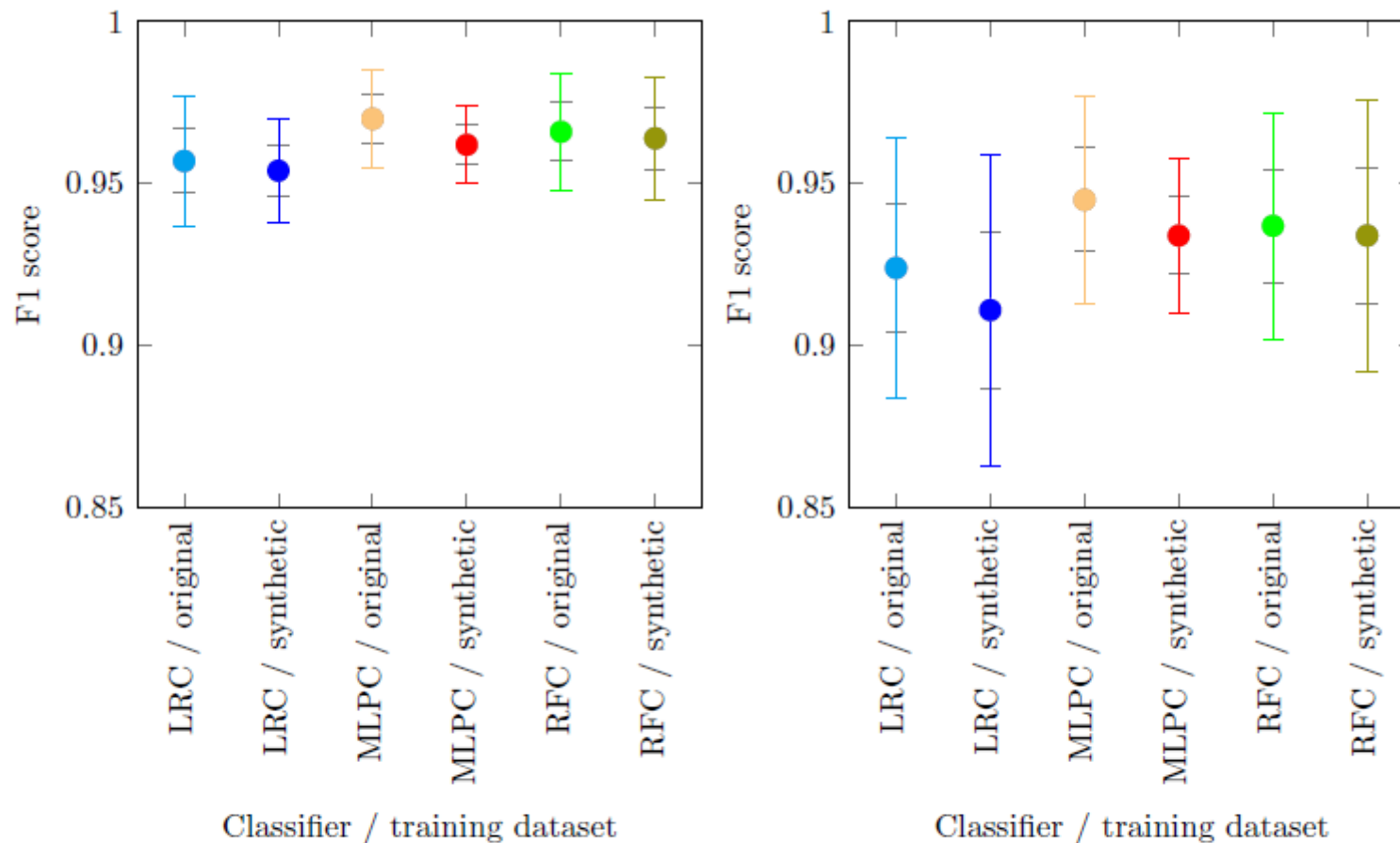


Figure 1: F1 scores for different classifiers trained on the original and synthetic datasets for Benign (left chart) and Malignant (right chart) classes with performance observed on the same original testing datasets. The dots indicate mean values and error bars indicate 1/2 and 1 standard deviations.

Example 2: Coimbra dataset

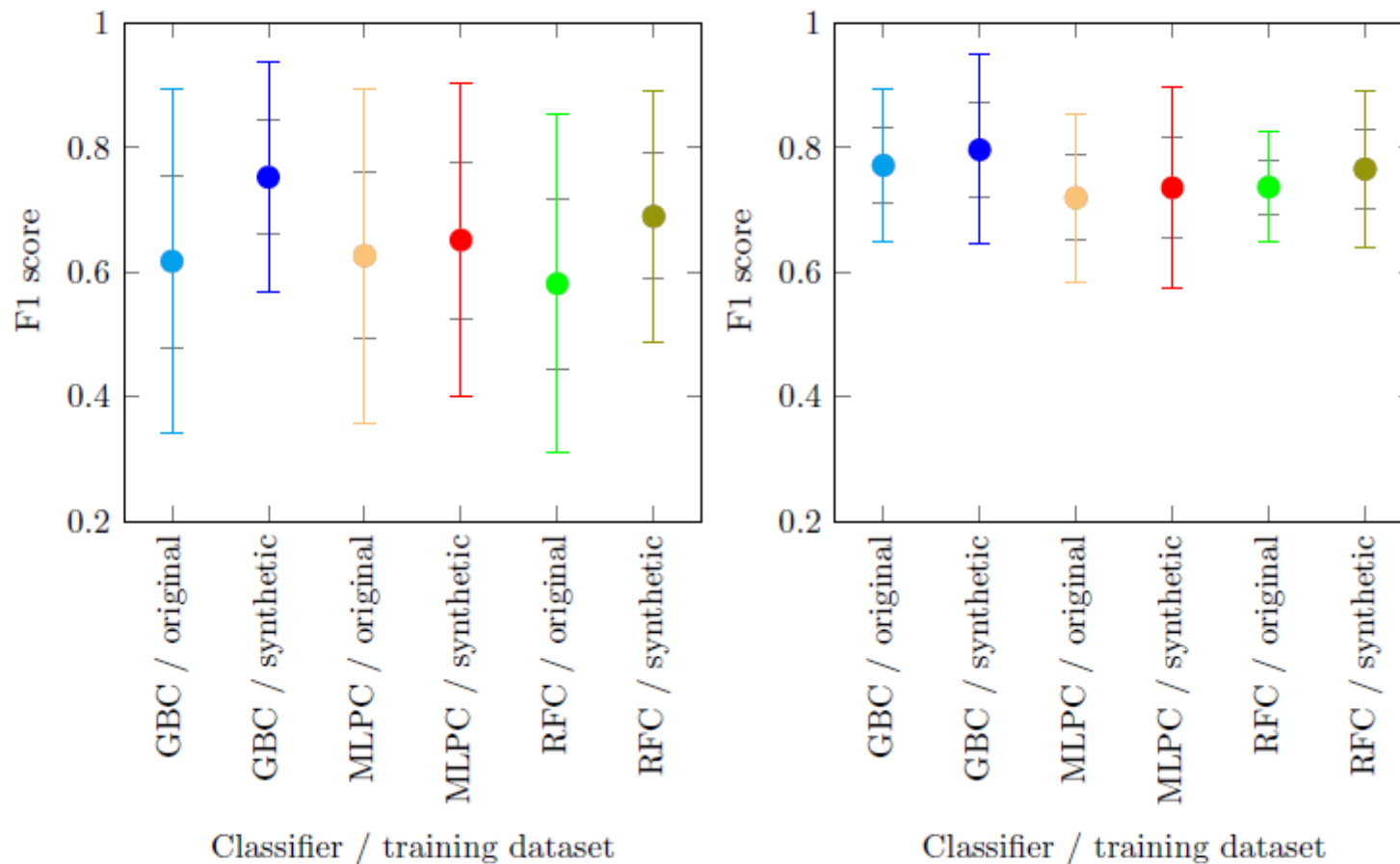


Figure 2: F1 scores for different classifiers trained on the original and synthetic datasets for the Healthy Control (left chart) and Patient (right chart) classes with the performance compared to the same original testing datasets. The dots indicate mean values and the error bars indicate 1/2 and 1 standard deviations.

Example 3: Australian credit approval dataset

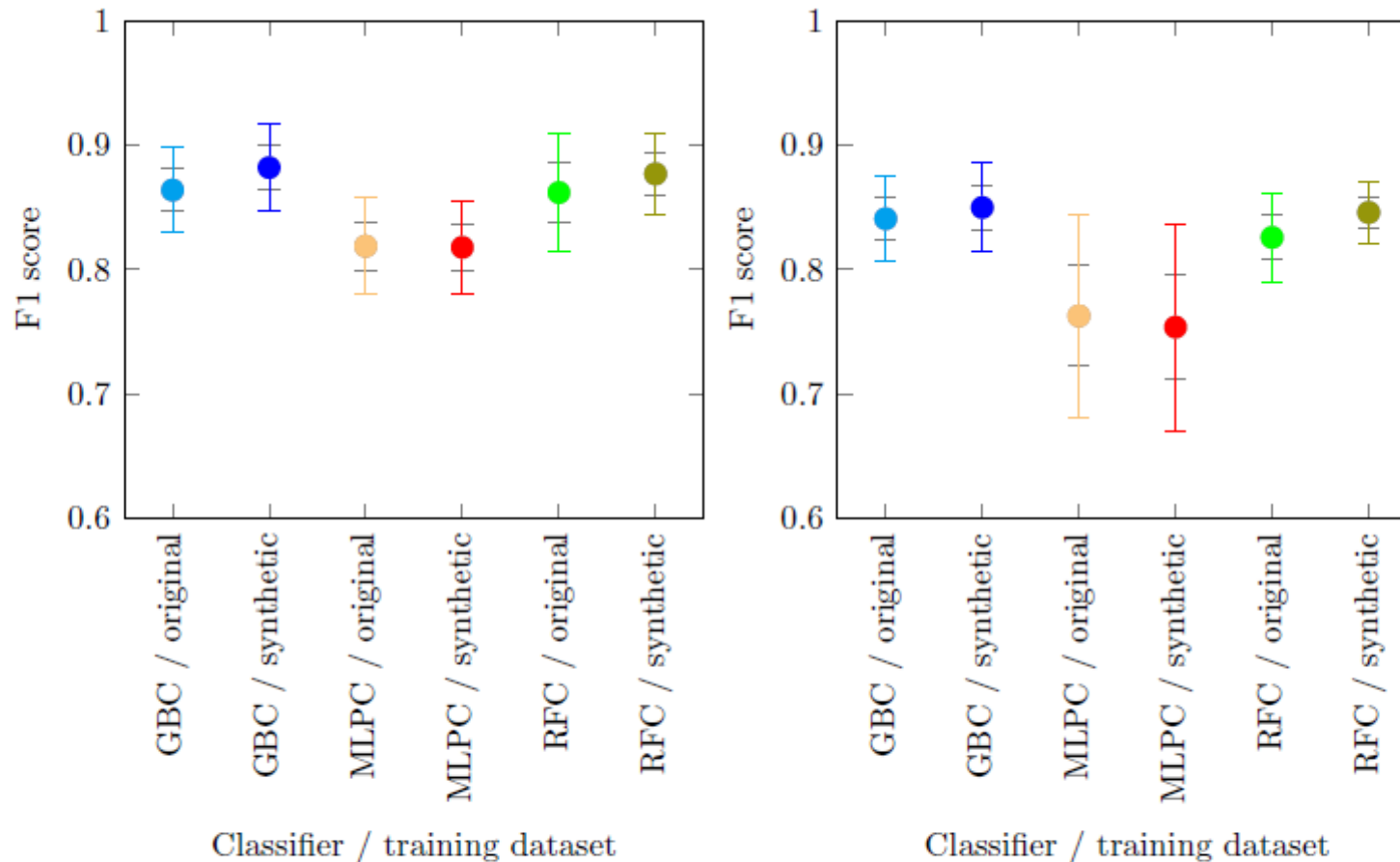


Figure 3: F1 scores for different classifiers trained on the original and synthetic datasets for Class 1 (left chart) and Class 2 (right chart), comparing the performance on the same original testing datasets. The dots indicate mean values and error bars indicate 1/2 and 1 standard deviations.

General Disclaimer

This communication has been prepared by Standard Chartered Bank. Standard Chartered Bank is incorporated in England with limited liability by Royal Charter 1853 Reference Number ZC18. The Principal Office of the Company is situated in England at 1 Basinghall Avenue, London, EC2V 5DD. Standard Chartered Bank is authorised by the Prudential Regulation Authority and regulated by the Financial Conduct Authority and Prudential Regulation Authority.

Banking activities may be carried out internationally by different Standard Chartered Bank branches, subsidiaries and its affiliates (collectively “SCB”) according to local regulatory requirements. With respect to any jurisdiction in which there is an SCB entity, this document is distributed in such jurisdiction by, and is attributable to, such local SCB entity. Recipients in any jurisdiction would contact the local SCB entity in relation to any matters arising from, or in connection with, this document. Not all products and services are provided by all SCB entities.

This material is provided for general information purposes only and does not constitute either an offer to sell or the solicitation of an offer to buy any security or any financial instrument or enter into any transaction or recommendation to acquire or dispose of any investment.

The information contained herein does not purport to identify or suggest all the risks (direct or indirect) that may be associated with conducting business.

This communication is prepared by personnel from SCB’s Sales and/or Trading Desks. It is not research material and is not a product of SCB’s Research.

Any views expressed may differ from those of SCB’s Research. This material has been produced for reference and is not independent research or a research recommendation and should therefore not be relied upon as such. The material has not been prepared in accordance with legal requirements designed to promote the independence of investment research. This material is not subject to any regulatory prohibition on dealing ahead of the dissemination of investment research.

This material is provided on a confidential basis and may not be reproduced, redistributed or transmitted, whole or in part, without the prior written consent of SCB.

This communication is not independent of SCB’s own trading strategies or positions. Therefore, it is possible, and you should assume, that SCB has a material interest in one or more of the financial instruments mentioned herein. If specific companies are mentioned in this communication, please note that SCB may at times seek to do business with the companies covered in this communication; hold a position in, or have economic exposure to, such companies; and/or invest in the financial products issued by these companies. This communication may be a marketing communication as referenced in the Market in Financial Instruments Directive 2014/54/EU.

SCB may be involved in activities such as dealing in, holding, acting as market makers or performing financial or advisory services in relation to any of the products referred to in this communication. The Sales and Trading personnel who prepared this material may be compensated in part based on trading activity. Accordingly, SCB may have a conflict of interest that could affect the objectivity of this communication.

This communication should not be construed as a recommendation (except to the extent it is an “investment recommendation” under MAR (as defined below)) for the purchase or sales of any security or financial instrument, or to enter into a transaction involving any instrument or trading strategy, or as an official confirmation or official valuation of any transaction mentioned herein.

The information provided is not intended to be used as a general guide to investing and does not constitute investment advice or as a source of any specific investment recommendations as it has not been prepared with regard to the specific investment objectives or financial situation of any particular person.

SCB does not provide, and has not provided, any investment advice or personal recommendation to you in relation to the transaction and/or any related securities described herein and is not responsible for providing or arranging for the provision of any general financial, strategic or specialist advice, including legal, regulatory, accounting, model auditing or taxation advice or services or any other services in relation to the transaction and/or any related securities described herein. The particular tax treatment of a service or transaction depends on the individual circumstances of each client and may be subject to change in the future. Accounting laws, rules, regulations, standards and other guidelines may differ in different countries and/or may change at any time without notice. SCB may not have the necessary licenses to provide services or offer products in all countries or such provision of services or offering of products may be subject to the regulatory requirements of each jurisdiction and you should check with your advisors before proceeding. Accordingly SCB is under no obligation to, and shall not determine the suitability for you of the transaction described herein.

You must ensure that you have sufficient knowledge, experience, sophistication and/or professional advice to make your own evaluation of the merits and risks of entering into such transaction. You are advised to make your own independent judgment (with the advice of your professional advisers as necessary) with respect to the risks and consequences of any matter contained herein.

General Disclaimer (continued)

While reasonable care has been taken in preparing this document, SCB expressly disclaims any liability and responsibility for any damage or loss you may suffer from your use of or reliance of the information contained herein. Any past or simulated past performance including back-testing, modelling or scenario analysis contained herein is not an indication of future performance or results. Changes in rates of exchange may have an adverse effect on the value of investments. No representation is made as to the accuracy of the assumptions made within, or completeness of, any modelling, scenario analysis or back-testing.

Predictions, projections or forecasts contained herein are not necessarily indicative of actual future events and are subject to change without notice. You are cautioned not to place undue reliance on such statements. While all reasonable care has been taken in preparing this communication, SCB makes no representation or warranty as to its accuracy or completeness.

Any opinions or views of third parties expressed in this material are those of the third parties identified, and not of SCB. Some of the information appearing herein may have been obtained from public sources and while SCB believes such information to be reliable, it has not been independently verified by SCB.

All opinions and estimates are given as of the date of the relevant document and are subject to change without notice. The value of any investment may also fluctuate as a result of market changes. SCB is not obliged to inform the recipients of this communication of any change to such opinions or estimates.

Where this material is an "investment recommendation" as defined in Article 3(1)(35) of the EU Market Abuse Regulation ("MAR"), distribution of this material is subject to the relevant provisions of MAR. For more information, please consult the MAR disclosures page.

Europe: This communication is not directed at Retail Clients in the European Economic Area as defined by Directive 2014/65/EU.

United States: Except for any documents relating to foreign exchange, rates or commodities, distribution of this document in the United States or to US persons is intended to be solely to major institutional investors as defined in Rule 15a-6(a)(2) under the US Securities Act of 1934. All US persons that receive this document by their acceptance thereof represent and agree that they are a major institutional investor and understand the risks involved in executing transactions in securities. Any US recipient of this document wanting additional information or to effect any transaction in any security or financial instrument mentioned herein, must do so by contacting a registered representative of Standard Chartered Securities (North America) LLC., 1095 Avenue of the Americas, New York, N.Y. 10036, US, tel +1 212 667 0700. WE DO NOT OFFER OR SELL SECURITIES TO U.S. PERSONS UNLESS EITHER (A) THOSE SECURITIES ARE REGISTERED FOR SALE WITH THE U.S. SECURITIES AND EXCHANGE COMMISSION AND WITH ALL APPROPRIATE U.S. STATE AUTHORITIES; OR (B) THE SECURITIES OR THE SPECIFIC TRANSACTION QUALIFY FOR AN EXEMPTION UNDER THE U.S. FEDERAL AND STATE SECURITIES LAWS NOR DO WE OFFER OR SELL SECURITIES TO U.S. PERSONS UNLESS (I) WE, OUR AFFILIATED COMPANY AND THE APPROPRIATE PERSONNEL ARE PROPERLY REGISTERED OR LICENSED TO CONDUCT BUSINESS; OR (II) WE, OUR AFFILIATED COMPANY AND THE APPROPRIATE PERSONNEL QUALIFY FOR EXEMPTIONS UNDER APPLICABLE U.S. FEDERAL AND STATE LAWS.

SCB DIFC: This communication is issued by Standard Chartered Bank, Dubai International Financial Centre Branch ("SCB DIFC"). SCB DIFC having its offices at Dubai International Financial Centre, Building 1, Gate Precinct, P.O. Box 999, Dubai, UAE is authorised by the Dubai Financial Services Authority ("DFSA"). This document is intended for use only by Professional Clients, as defined by the DFSA Rulebook.

For SCB DIFC Islamic business: For Islamic banking business, SCB DIFC act under the supervision of the banks Shari'ah Supervisory Committee, which currently has three members. Relevant information on the Shari'a Supervisory Committee for the time being is available on the Standard Chartered Bank website in the Islamic banking section (at the time of publication of this Country Supplement).

Upon your request we will provide details of the manner and frequency of Our Shari'a reviews. The special features and characteristics of the Products we offer are set out in the Shari'a Compliant Agreements. We do not commingle funds attributable to our Islamic finance business with funds attributable to conventional finance business, by ensuring segregation in our booking systems by various system tools such as separate product codes, booking codes, branch codes as appropriate.

© 2018 Copyright Standard Chartered Bank and its affiliates. All rights reserved. All copyrights subsisting and arising out of these materials belong to Standard Chartered Bank and may not be reproduced, distributed, amended, modified, adapted, transmitted in any form, or translated in any way without the prior written consent of Standard Chartered Bank.