Quantitative Monographs
Introduction to Deep Learning

Equities

Global
Quantitative

David Jessop
Analyst
david.jessop@ubs.com
+44-20-7567 9882

Oliver Antrobus, CFA
Analyst
oliver.antrobus@ubs.com
+61-3-9242 6467

Luke Brown, CFA
Analyst
luke.brown@ubs.com
+61-2-9324 3620

Cathy Fang, PhD
Associate
cathy.fang@ubssecurities.com
+86-213-866 8891

Josie Gerken, PhD
Analyst
josephine.gerken@ubs.com
+44-20-7568 3560

Josh Holcroft
Analyst
josh.holcroft@ubs.com
+852-2971 7705

Desi Ivanova
Associate Analyst
desi-r.ivanova@ubs.com
+44-20-7568 1754

Claire Jones, CFA
Analyst
claire-c.jones@ubs.com
+44-20-7568 1873

Pieter Stoltz
Analyst
pieter.stoltz@ubs.com
+61-2-9324 3779

Paul Winter
Analyst
paul-j.winter@ubs.com
+61-2-9324 2080

Shanle Wu, PhD
Analyst
shanle.wu@ubs.com
+852-2971 7513

Expert call on machine learning

We are happy to have hosted a call with Matthew Dixon, Professor of Finance & Statistics at the Illinois Institute of Technology on the topic of deep learning in finance. This note is our edited transcript of the call where we have added in the contents of the slides.

Introducing deep learning

The key idea of deep learning is to build a multi-layer network and these hierarchical layers represent a very high dimensional input space. Deep learning doesn't make sense when you only have a small amount of input data. Even daily stock return data might not be enough.

Introduction[1]

This is a presentation on deep learning in finance. The main takeaway from this talk is going to be identifying and understanding where deep learning makes sense to apply within finance and, in some sense, that means untying some of the rhetoric which has been coming from Silicon Valley around this topic.

Machine learning is an algorithmic class (Breiman, 2001) of models that differ from the classical statistics models in that no assumptions about the data generating process are made. Unlike, for example, Brownian motion and other typical models that have come to be celebrated within finance, this class of model fitting or *learning* does not assume any (distributional) properties of the data that it is fitted to, thus giving rise to a much wider class of modelling capacity.

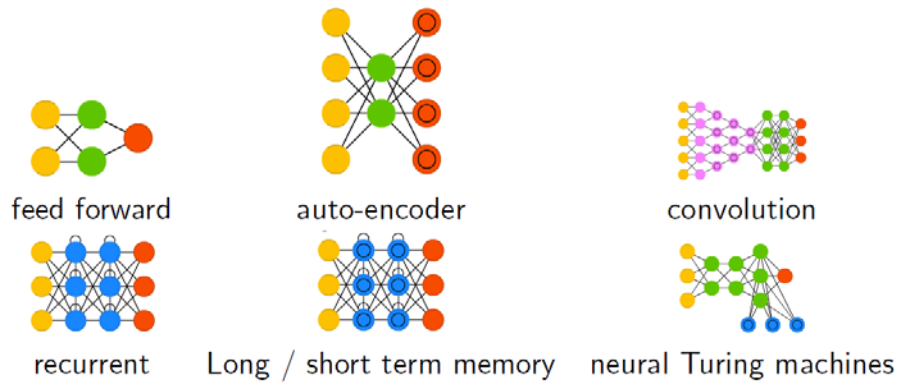
Deep learning requires lots of input variables

The key idea behind deep learning is the use of hierarchical layers to represent (usually) a very high-dimensional *input* space[2]. It really doesn't make sense to use deep learning if you have a small number of input variables (or *features*). I'm going to discuss and provide a little insight into how deep learning works as opposed to just presenting performance results (which will also be done towards the end of the presentation).

Traditional metrics fail in high dimensions

We're going to depart a little from the traditional statistical metrics such as R^2, *p*-values, *t*-values and the concept of statistical significance overall. These often fall short as a measure of out-of-sample forecasting accuracy, particularly for high dimensional problems. Instead we turn to a well-worn concept within pattern recognition and machine learning known as the *bias-variance trade-off*. The idea behind it is to control how well the model performs out-of-sample and also to control the difference in accuracy of the model in-sample and out-of-sample.

Figure 1: Deep Architectures in TensorFlow



Source: Matthew Dixon. Some of the most commonly used deep learning architectures for modelling.

Machine learning is good at avoiding overfitting

And that's really what machine learning does well – it enables us, through a framework which is based on information theory, to *avoid overfitting*. Of course, it does require some skill; there are some techniques that we'll just touch on during

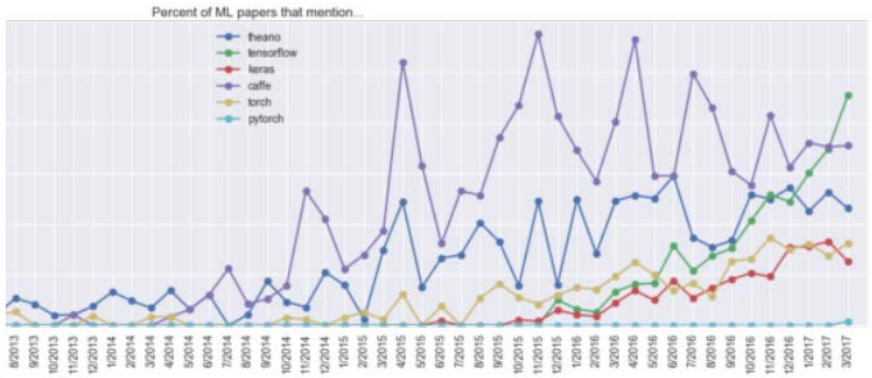
[1] This is our edited transcript of the Expert Call with Professor Matthew Dixon. Produced with his permission. We have edited some of the slide content and equations into the text for clarity.
[2] Input space is the set of all possible realisations of the input variables.

this presentation and as we go along I'll leave some pointers for anyone who is interested in learning more.

Why I'm giving this talk in part is because we've seen a surge of interest in machine learning algorithms (partly driven by the Googlization trend) and a number of these include various neural network architectures (some of which are illustrated in Figure 1). A lot of these architectures are now very much in place and used widely in practice: whether it's speech recognition on your phone; whether it's image recognition; whether it's self-driving cars. However, there are certain subsets of these architectures which are more relevant to finance, such as recurrent neural networks, which can handle time series data; more standard architectures (e.g. feed forward) might also be useful. All of these techniques can be regarded as a form of data compression, which is a way of taking fairly noisy data and compressing it down into several hundred parameters.

Google's TensorFlow, which is a software library for machine learning, has been growing rapidly (see Figure 2) and a number of deep learning researchers are reporting to be using this. We are seeing an age where the volume of data is large and it's also cheap to do large computation, which allows us to be a little bit more relaxed about our data assumptions. That's what machine learning really is: it is us bypassing any parametric assumptions for the data and instead jumping forth with all our heavy machinery, taking advantage of our multi-core computing era that we are now in.

Figure 2: Growth of TensorFlow



% of papers	framework	has been around for (months)
9.1	tensorflow	16
7.1	caffe	37
4.6	theano	54
3.3	torch	37
2.5	keras	19
1.7	matconvnet	26
1.2	lasagne	23
0.5	chainer	16
0.3	mxnet	17
0.3	cntk	13
0.2	pytorch	1
0.1	deeplearning4j	14

Source : Andrej Karpathy's arXiv-sanity database

Source: Matthew Dixon

Examples using an unfair coin

I'm going to use a coin tossing as a way to introduce some of the theory and terminology that we're going to use for deep learning. Of course, many will be familiar with the main ideas of machine learning, but it doesn't hurt to tie this back to probability theory in an intuitive way.

Let's walk through a couple of examples slowly building up the complexity before we get to machine learning, deep learning, and some of the key concepts that are

crucial to learning in general. Then I'll talk more about the architecture side and application within equity research.

Running example: coin tossing experiment

The running example through this presentation will be an experiment of tossing a coin that shows heads with probability $p \in [0,1]$. We introduce a random variable Y, defined on a sample space, which contains a head or a tail, $\Omega = \{H, T\}$. We therefore have that Y is a Bernoulli random variable with probability mass function, f, so that:

$$f(Y=1) = P\{H\} = p \;\; and \;\; f(Y=0) = P\{T\} = 1-p$$

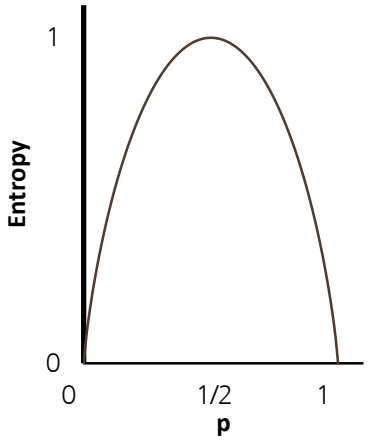
Entropy is a key concept underlying machine learning

The key concept underlying machine learning is the concept of entropy brought to us by Claude Shannon in his 1948 Bell System Journal paper. The idea is very simple: we want to understand how many bits of information we need. In this case we have a coin – which is a simple piece of binary information and we want to encode how much uncertainly there is about the experiment. The entropy of the binary variable Y under f is:

$$\mathcal{H}(f) = -p\log_2 p \; -(1-p)\log_2(1-p) \leq 1bit \; [3] \qquad (1)$$

Figure 3 plots the entropy of Y as a function of the probability of getting a head, p, showing that the entropy (or uncertainy) is maximum when we have a fair coin and decreases to zero when the coin always shows either a head or a tail.

Figure 3: Entropy as a measure of uncertainty



Source: UBS Quant. The figure illustrates the entropy of a binary random variable.

For an arbitrary density function f (i.e. not necessary a Bernoulli mass function), the entropy is defined as minus the expected value of the log density:

$$\mathcal{H}(f) = -\mathbb{E}_f(log_2 f).$$

In the simple case of a Bernoulli variable, it is easy to see that we recover our definition of binary entropy (see Equation 1), which, of course, is less than or equal to one bit (see Figure 3). In the general case of N states the entropy is less than or equal to $\log_2 N$.

We consider our random variable more formally as a map from the sample space Ω to some subset of the real line, $Y: \Omega \to \Upsilon \subset \mathbb{R}$ (although Y will only take on discrete values). We're going to assume that there's a density function for this random variable, which is parameterised by a parameter θ: $f(y|\theta) = P(\omega \in \Omega: \Upsilon(\omega) = y)$. The meaning of the parameterization will become very clear in a moment, when we look at machine learning, because the goal of machine learning is essentially to learn a set of parameters.

Machine learning is really about minimising cross-entropy

Cross-entropy is the last key concept we need to introduce, which will allow us to quantify how different two probability distributions are. This is important in machine learning because our aim is to approximate the true (but unknown) distribution $f(y|\theta)$ with something we can estimate, $f' = f(y|\theta')$. More formally, we wish to minimise the cross-entropy between two probability distributions f and f', which is defined as the negative expected value of the logarithm of $f' = f(y|\theta')$, where the expectation is taken under the true distribution f:

$$\mathcal{H}(f,f') = -E_f[log_2 f'] \; = -\sum_y f(y|\theta) log_2 f(y|\theta') \geq \; \mathcal{H}(f)$$

[3] Entropy is usually measured in *bits* – this is the case when log-base 2 is used. One can also use a different base; common examples include the natural log (units of entropy are then called *nats*) or log base 10 (*bans*).

Notice that the cross-entropy of f with itself is simply the entropy: $\mathcal{H}(f,f) = \mathbb{E}_f(log_2 f) = \mathcal{H}(f)$.

Going back to our unfair coin example, we estimate $f(Y=1|\theta') = p_{\theta'}$ so that it matches the underlying true distribution, $f = f(y|\theta)$ as closely as possible. To this end, given a set of realisations $\boldsymbol{y} = \{y_1, \ldots, y_n\}$ of our random variable Y, we "learn" the best model $\hat{f} \coloneqq f(y|\hat{\theta})$ which minimises the cross-entropy:

$$\hat{\theta} \coloneqq \underset{\theta'}{\operatorname{argmin}}\, \mathcal{H}(f,f')$$

Log-likelihood as approximation of cross-entropy

It is very common in applications, and in particular for a lot of financial and econometrics modelling, to assume that the log-likelihood is a good approximation of the negative cross entropy[4]:

$$-\mathcal{H}(f,f') \approx l(\theta'|\boldsymbol{y}) = \frac{c}{n}\sum_{i=1}^{n} \ln f(y_i|\theta')$$

That's one of the reasons why the log-likelihood function is so heavily used in finance – for example, maximum likelihood methods are commonly used to estimate the parameters in various geometric Brownian motion models.

Back to our coin tossing experiment, suppose we have a sequence of ones and zeros, each of which is a realisation of Bernoulli trials (e.g. **y** = {1, 0, 1, 1, 1, 0, 1, 0}), then the log-likelihood becomes:

$$l(\theta|\boldsymbol{y}) = \ln\left(p_\theta^{y_1}(1-p_\theta)^{y_1} \times \ldots \times p_\theta^{y_n}(1-p_\theta)^{y_n}\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} y_i \ln p_\theta + (1-y_i)\ln(1-p_\theta)$$

Vanilla maximum likelihood often results in overfitting…

That is where regularisation comes into play

In many circumstances however the maximum likelihood approach is not a good approximation of the cross-entropy. Another problem associated with maximum likelihood is that it often leads to overfitting. Various **regularisation techniques** have been developed to try and overcome some of these problems, and this continues to be an active area of machine learning research. A simple example of regularisation is the well-known Akaike Information Criteria (AIC, a form of L2 regularisation), introduced in the 1970s.

Before we jump into machine learning, it is helpful to build up our coin tossing experiment one more level. This time we assume that we have two coins and let X and Y be two random variables that count the number of heads generated by each coin. Furthermore, the two coins are correlated[5] and we assume that there is conditional density function for Y given X: $f(y|x,\theta)$ with $f(y=1|x,\theta) = p_\theta(x)$ In other words, the number of heads generated by one coin is dependent on the number of heads generated by the other coin. This stylised example of correlated coins is useful for machine learning as it illustrates the concept of observing one variable and then making a statement about the other one.

Same as before, we can write the cross-entropy between the true density and some approximation of it, $f' \coloneqq f(y|x,\theta')$ as:

[4] c is a constant equal to the conversion from natural log to log base 2

[5] The UBS Quant Team has discussed how you can create correlated coins and we have to admit we've failed to come up with a suitable physical mechanism. It still works as a thought experiment though.

$$\mathcal{H}(f(\cdot\,|x), f(\cdot\,|x,\theta') = -p(x)\log_2 p_{\theta'}(x) - \big(1 - px(x)\big)\log_2(1 - p_{\theta'}(x)).$$

Given a set of realisations of X and Y, $D = \{x_i, y_i\}_{i=1,\dots n}$, the likelihood then becomes $l(\theta|D) = \frac{1}{n}\sum_{i=1}^{n} \ln f(y_i|x_i,\theta')$ and is maximised for a choice of $\theta = \hat{\theta}$.

What is machine learning…?

Machine learning is all about finding the map between X and Y, where X, does not need to be the outcome of a coin, but some real input space, which we assume to be high dimensional. As a simple analogy one can think of the outcome of one coin as being dependent on hundreds or thousands of other coins. What we are then trying to do is to construct a nonlinear predictor, $\hat{Y}(X)$, of an output, Y given a high dimensional input matrix[6] X. Put differently, machine learning is just the study and the construction of an input/output map.

… A map from a high dimensional input space to an output space

The output variable Y can be continuous, discrete or mixed. If Y is discrete then we call this a *classification problem*; otherwise we call it a *regression problem*. We'll see later that in simple cases, this map reduces to an ordinary linear regression or a logistic regression (a type of semi-affine transformation).

Deep predictors use a sequence of hidden layers

What is a deep predictor? A deep predictor is trying to approximate this map between input data and some output using a sequence of *hidden layers*. We can assume that there are L layers and there's now a composition of some multivariate functions between the layers taking our input data and transforming them through some semi-affine function and composing them together to give us our sequence of layers.

Definition – Deep Predictor

A deep predictor is a particular class of multivariate function F(X) constructed using a sequence of L layers via a composite map

$$\hat{Y}(X) \coloneqq F_{W,b}(X) = \big(f^{L}_{W^L,b^L} \circ \cdots \circ f^{1}_{W^1,b^1}\big)(X)$$

$f^{l}_{W^l,b^l} \coloneqq f^{l}(W^l X + b^l)$ is a semi-affine function, where f^l is univariate and continuous.

$W = (W^1, \dots, W^L)$ and $b = (b^1, \dots, b^L)$ are weight matrices and offsets respectively.

The theory behind deep learning is now well understood

"bio analogy" is no longer necessary

There is a lot of theory behind deep learning and it's well understood why applying such transformations to the input space makes sense and works in practice (i.e. it's not a purely empirical result that has seen great success in applications). What we need to do is learn from observed data the weights, W, and biases, b, that link the nodes in the neural network from one layer to another. The analogy with neurons - the "bio analogy" - may have served us well in the past, but isn't needed now. I think we understand deep learning enough in mathematical terms that we needn't rely on biological analogies to explain things, which I think comes with some relief to those that have given up on genetic algorithms and that whole area of bio-inspired algorithms. There's no doubt in my mind that, really, these neurons are just transformations with some semi-affine function (whether the brain itself does that is questionable).

The structure of a deep prediction rule can be written as a hierarchy of $L-1$ unobserved hidden layers (and a final output layer, giving us a total of L layers), $\{Z^l\}_{l=1,\dots,L-1}$, given by:

[6] We are now using X as the data matrix, not as a random variable.

$$\hat{Y}(X) = f^L(Z^{L-1})$$
$$Z^0 = X,$$
$$Z^1 = f^1(W^1 Z^0 + b^1),$$
$$\ldots$$
$$Z^{L-1} = f^{L-1}(W^{L-1} Z^{L-2} + b^{L-1})$$

The functions $f(.)$ are known as *activation functions*. Depending on the application (classification vs regression), popular choices include rectified linear unit ($ReLU$), hyperbolic tangent ($\tanh(.)$), max and softmax (usually used in the final/output layer in classification problems).

There are a number of questions that come to mind: How many layers do I need in my application? How do I fit these weights and biases? The answers to such questions still remain a challenge in machine learning; however the availability of software, such as TensorFlow, provides us with a very efficient way to solve some of these problems.

Why does deep learning work? Let us consider the two-dimensional space as shown in Figure 4. We have two lines labelled one for the red line and zero for the blue. These are linked to coordinates in this two-dimensional space.

We want to separate the input space in order to distinguish between the blue line and the red line. If we were going to use a linear method such as logistic regression as shown in Figure 5, we would struggle to accurately classify these two lines.

Figure 4: Problem

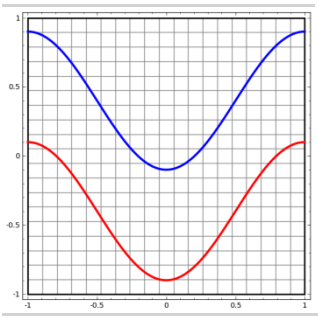


Figure 5: Linear method

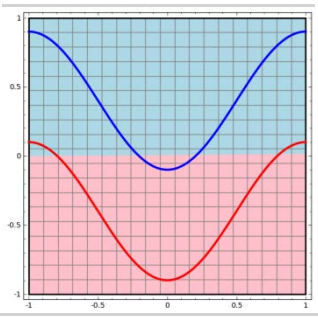


Figure 6: Transformation

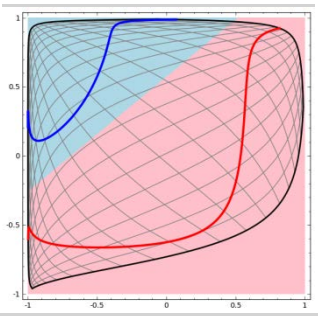
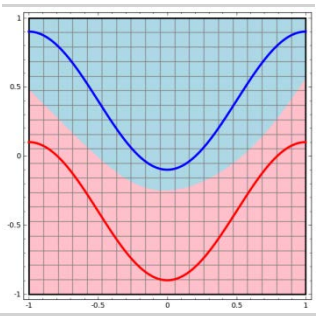


Figure 7: Final result



Source: Chris Olah. Used with permission

What do hidden layers do?

What a hidden layer would do is apply a semi-affine function, transforming the input space into something that the output layer could use. As Figure 6 demonstrates, by doing this we are able to transform the input space so that it is easy for a linear separator to clearly discern between the red and the blue line. Translating back to the original space (Figure 7) we see that the boundary is no longer linear. The more layers we have, the more complex (and non-linear) a transformation you tend to do, especially when used in a much higher dimensional space.

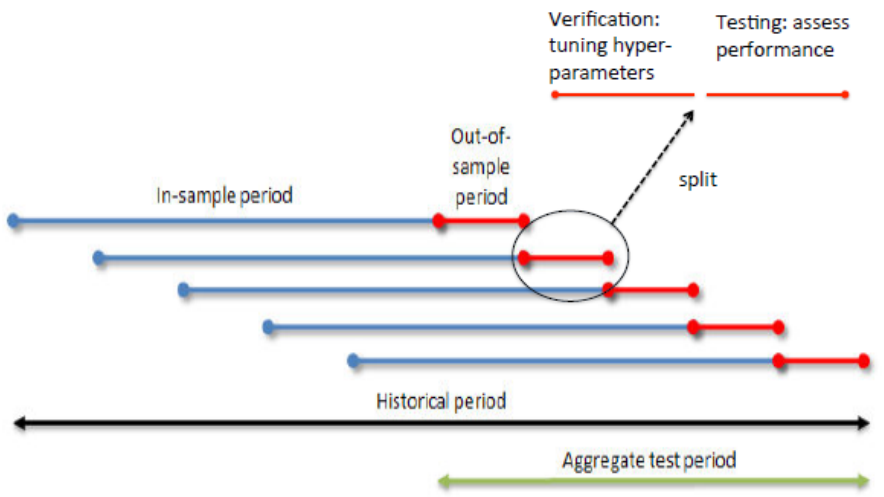
This simple example illustrates the role of hidden layers and therefore the main ideas of deep learning. Of course, different learning methods perform different types of transformations. There is still a lot of work to be done on exactly how these are done and in particular what types of transformations would be most useful for noisy financial data, which in most cases comes in the form of time series. What is special about time series modelling is that we can't simply apply the standard methods for choosing appropriate parameter values, such as (K-fold)

cross-validation or bootstrapping, as those will break the temporal structure of the data.

Cross validation for training time-series models

To train a time series model we are going to use a technique known as *time series cross-validation*, which is demonstrated in Figure 8. We will assume there's a period of historical data and on a rolling (or expanding) window basis, we use part of this historical data for **in-sample training** of out model; *this is known as the training set*. We have an out-of-sample period which the model has not seen that will be used for **tuning** the parameters of the neural network; *this is known as the validation set*. Ideally, we would like to keep some of the out-of-sample period datapoints to **test** the final (tuned) model; *this is known as the test set*. This will give us an unbiased view of how accurate the model will perform in the future (as the tuned architecture hasn't seen it before).

Figure 8: Time series cross validation



Source: Matthew Dixon, used with permission. Times series cross validation, also referred to as walk forward optimization, is used to avoid introducing look-ahead bias into the model.

We'll keep sliding the window forward. As with optimization there are variants on this. This approach I think is common to us and seen before in finance, although I have seen still an alarming number of research papers reporting in-sample model performance and not assessing over the moving time frame what is the real performance of a strategy.

Model training

Given some data, $D = \{Y_i, X_i\}_{i=1,\ldots,n}$ the first part of the process – model training, is simple – it is simply an optimisation problem that involves minimising some loss function $\mathcal{L}$, solving for the weights W and bias terms b:

$$\underset{\mathrm{W,b}}{\operatorname{argmin}} \sum_i \mathcal{L}\left(Y_i, \hat{Y}^{W,b}(X_i)\right)$$

When Y is categorical, we will use the negative log-likelihood as approximation of the cross-entropy.

$$\mathcal{L}\left(Y_i, \hat{Y}^{W,b}(X_i)\right) = -y_i \log \hat{Y}^{W,b}(X_i) + \phi(W, b)$$

When Y is continuous a popular choice of the loss function is the $L2$ norm an error metric, in which case the problems simplify to traditional least squares (with a penalisation term):

$$\mathcal{L}(Y_i, \hat{Y}^{W,b}(X_i)) = \left|\left|Y_i - \log \hat{Y}^{W,b}(X_i)\right|\right|_2^2 + \phi(W, b)$$

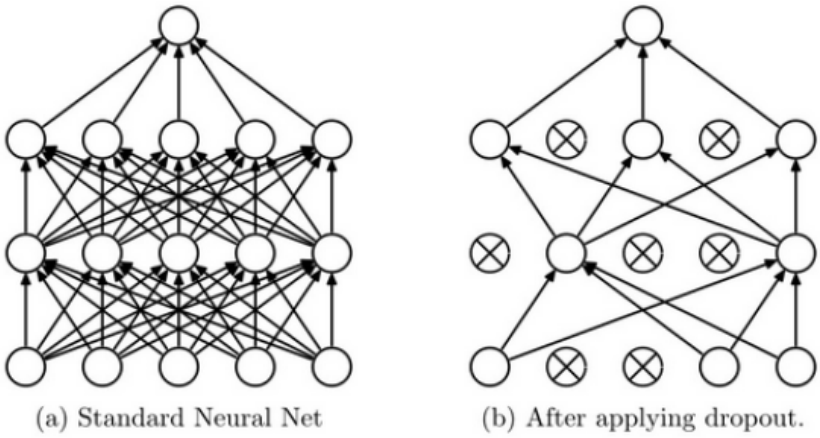
TensorFlow to estimate the weights and bias terms

Regardless of the choice of a loss function, to be able to find a minimum, we need to compute its gradient and this is where make use of the power of Google TensorFlow.

Dropout: regularisation technique to prevent overfitting

I'll tie that in with a technique which is called *dropout*, illustrated in Figure 9, where we're now trying to take an average of sub-dimensional models. We're dropping out some of the nodes, we're measuring accuracy of the model, and essentially we're repeating this over and over again like an ensemble approach with some reduced model and trying to assess the performance under various different randomly eliminated nodes. This is a technique which has been very effective in reducing the dimension of the network and variable selection, and it leads to improved performance out of sample.

Figure 9: Dropout



Source: Matthew Dixon. Used with permission.

Dropout can be combined with stochastic type of search:

$$D_i^l \sim Ber(\theta)$$
$$\tilde{Z}^l = D^l \times Z^l, \quad 1 \leq l \leq L$$
$$Z^l = f^l(W^l \tilde{Z}^{l-1} + b^l) \,,$$

and can be thought of as a ridge regression with a g-prior. In other words, there is a Bayesian of interpretation to this technique, details of which can be found in a paper by Polson and Sokolov (2017)[7].

An example: portfolio returns

Predicting the direction of the return of a portfolio

As an example we are going to consider a portfolio with a position in n stocks. We want to make an investment decision every day based on using a large number of stocks over history. We are going to attempt to apply deep learning to the problem capturing directional changes in the portfolio returns. Specifically, we're trying to make a decision on whether to long or short the entire portfolio based on our prediction of the next-day portfolio return.

[7] Polson, N & Sokolov V. (2017). Deep Learning: A Bayesian Perspective. Available at https://arxiv.org/abs/1706.00473

We calculate the portfolio returns using equal weighting. The data we are going to use is daily data. What we want to do is use the deep network to maximize the capacity to predict the next day portfolio returns given our cross-sectional market data. So, given the returns across all the stocks, what can we say about the overall portfolio return in the future?

Deep learning is very data hungry

It is worth stopping to mention here, with this notion of horses for courses, deep learning does very well when there's a high-dimensional input space and there's a lot of data, a lot of historical data. Now, we're stretching it a bit with daily data, and I've actually been working a lot more with intraday data and the reality is that deep learning is very data hungry and to avoid over-fitting we really need to train on the large samples.

It doesn't make sense to apply deep learning as a predictor where we might use an autoregressive, ARIMA or GARCH-type model; a univariate type of model. It makes sense to use it in a highly multi-variate model, in this case a classification model based on a large number of inputs. And large here is hundreds, ideally thousands or tens of thousands.

We define our stock returns as $r_i(t)$ and hence our portfolio return is

$$r_P(t) = \sum_{i=1}^{n} w_{i,t} r_i(t), \quad w_{i,t} = \frac{1}{n}$$

The observed data consists of the historical return of all the stocks, and then we are going to label the data with a categorical variable

$$Y_t = \begin{cases} 1, & r_p(t+1) > \epsilon \\ 0, & |r_p(t+1)| \leq \epsilon \\ -1, & r_p(t+1) < -\epsilon \end{cases}$$

The goal will be to learn the map $\hat{Y}_t = F_{W,b}(X_t)$. ϵ is a threshold chosen from the training data to avoid an imbalance of the labels.

Hence we are going to design a simple strategy $S(Y_t)$ which choses whether to hold a long, neutral or short position in all the stocks over the next period.

$$w_{i,t+1} = S(Y_t) = \begin{cases} \frac{1}{n}, & Y_t = 1 \\ 0, & Y_t = 0 \\ -\frac{1}{n}, & Y_t = -1 \end{cases}$$

This example is trying to exploit the high dimensions in our data set.

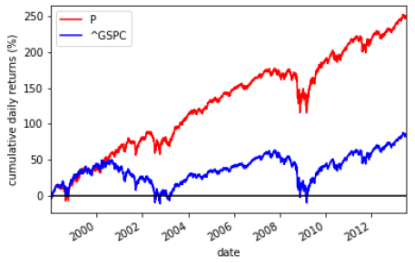
The returns of the portfolio, P, are shown in Figure 10. It's a subset of around 385 stocks from the S&P where we have a history over a 15-year period.

The experiment is set up to use a training horizon of 3,840 days, a test horizon of 30 days and then retraining the model every 30 days over 300 test observations. We avoid look-ahead bias by normalizing the data based on means and standard deviations from the training data. This is a common mistake we often see as we tend to normalize based on the full data, not just on the training data. If we use the full data set then we have a look-ahead bias.

In order to benchmark our deep learning model we will use logistic regression with regularization (to deal with multicollinearity) as a baseline classifier. Our neural network has three hidden layers with a "tapered" feed-forward architecture (with 200, 100, 50 neurons in each layer). This is a well-known successful configuration. This example does not consider how to capture auto-correlation within the data; it is cross-sectional only.

We show the performance of the logistic regression in Figure 11 and the neural network in Figure 12. We use a number of common measures of performance: precision, recall and the f1-score.

Figure 10: Portfolio return



Source: Matthew Dixon, used with permission.

Results

Figure 11: Performance of logistic regression

	in-sample				out-of-sample			
label	precision	recall	f1-score	support	precision	recall	f1-score	support
-1	0.91	0.88	0.89	1222	0.26	0.32	0.29	82
0	0.90	0.90	0.90	1347	0.38	0.35	0.36	121
1	0.89	0.92	0.91	1271	0.37	0.34	0.35	97
avg / total	0.90	0.90	0.90	3840	0.34	0.34	0.34	300

Source: Matthew Dixon, used with permission. The bias-variance trade-off is characterized by the difference between the in-sample and out-of-sample performance.

We can compare the in-sample logistic regression performance with the out-of-sample and we see a very big drop off. This is a very large variance and is very common with noisy data in finance. Our baseline is to beat this out-of-sample performance, the 0.34 for both the recall and the f1-score.

Figure 12: Performance of deep neural networks

	in-sample				out-of-sample			
label	precision	recall	f1-score	support	precision	recall	f1-score	support
-1	0.94	0.95	0.94	1222	0.35	0.35	0.35	82
0	0.97	0.93	0.95	1347	0.44	0.52	0.48	121
1	0.93	0.94	0.93	1271	0.41	0.32	0.36	97
avg / total	0.95	0.94	0.94	3840	0.41	0.41	0.41	300

Source: Matthew Dixon, used with permission. The bias-variance trade-off is characterized by the difference between the in-sample and out-of-sample performance.

When we look at the deep network, it is performing better in-sample than the logistic regression. More importantly, it's performing better out-of-sample. There is still a very big difference. It could be that we don't have enough data to fully represent our problem; it could be that we haven't got enough variables that is or enough training periods and there are more diagnostics than needed here. The deep network is also outperforming the simple logistic regression.

We're retraining the model every 30 days. You could of course train more frequently to improve performance more, especially if there's been some market

dislocation. It seems to me that we can certainly do better and outperform it and also reduce our risk. And certainly there's no question of the deep network outperforming a linear method. Several experiments show how well it can perform against other machine learning and so it does better there, too.

In passing that these types of approaches are very useful for high frequency trading, training the model on the data of the limit order book; there's a lot of information there.

In summary, deep learning is a form of machine learning for non-linear high-dimensional data. Regularization is really the key part here. The link between the log likelihood and the cross entropy, and that is really essential to how the learning approach works. I think there are many interesting applications to explore predicting the returns of portfolios, EFTs, where we are trying to model the entire all but constituent components as opposed to the aggregate result.

Clearly with so much information available there, why discard it? And there are obviously many other sources of data from exchange fundamentals and in equities and of course we could look across multiple exchanges and there are of course textual data in the form of news that can be encoded into this high-dimensional input space.

Questions[8]

How do you deal with liquidity? Or how do you model the impact of your trades on both prices and other people in the market?

I think my answer to that is you have to separate out the prediction from the rest of the problem. So make your prediction and then analyse how this translates into P&L.

My approach is to separate those out and to build your highly-predictive classifier and then to assess the impact of classification error under various different probabilities of fills, various other sort of trade fill constraints so that it becomes clear what the implications of error in the model are on the P&L.

There are other approaches which have tried to combine everything together. Reinforcement learning is an example. Michael Kearns at the University of Penn State wrote a paper about this some time ago. The outcome of the model is the trading position. There you can incorporate liquidity considerations and so on. That's not proven in any research to have worked well or any better than separating out the problem.

My bet is on classifying first with the data, taking a predictive signal and then having a separate model which then tries to assess the effect of liquidity constraints etc. This could be by back testing, or running a simulation of the market.

How does the size / frequency of the data effect the calculation time? Does this increase the latency of the algorithm?

There is a misconception around the implementation of these types of algorithms and the practicality of their deployment in a low-latency environment, or an environment where latency is a critical factor.

It is the training of the model which is computationally intensive, and that needn't be done on that frequently. I've done studies where I found surprisingly once I trained the model—this is from micro structural data—even with events that disrupt the market, the model remains fairly stable and does not to be retrained that often. I've seen the same also looking at daily data. Retraining is also computationally expensive. We would never want to do that in real time or near real time.

You can encode the trained network in C++ for example, and this can run the very fast.

I think the real problem is to have enough data to avoid over-fitting, because with machine learning one is trying not to impose a data generation process. If you do impose one, then this reduces the amount of data you require: with a parametric distribution you need less data.

When you relax all of that you need much more data – and that is the main challenge within finance. The latency is a much smaller problem. A network can be very efficiently implemented, especially with all the expertise that we have from the HFT quant community.

[8] Here we have edited the questions and answers quite significantly, but tried to keep the important points that were made.

How should one normalise the data when it isn't stationary? Should one make it stationary?

In the example above I'm not assuming any stationarity. The data is being normalised simply to ensure that the data is well-scaled. Deep Learning networks are remarkably resilient to different choices of scaling. The network and some of the algorithms which are performing the back propagation are quite robust.

What choice of network did you fit to the data? Did you use long / short memory networks or recurrent neural networks as the inputs are time series?

It's possible to actually encode lagged input variables into a feed-forward architecture. We have say P input variables and then you lag N times, you would then end up with P times N input variables. This then creates a much more difficult problem to solve because you need in a feed-forward architecture with many more neurons.

Of course if you use recurrent neural networks (RNNs), the more natural choice for capturing memory, then this reduces the number of neurons you need. You don't need to embed lagged variables in your input space. Hence your input space is a little smaller and generally you need fewer neurons and fewer hidden layers.

When I've experimented with long/short-term memory networks (LSTM), I've not seen any gain from using them. The idea is you would capture longer-term market movements and shorter-term movements. They are motivated in by work in say text completion where you're trying to close in a parenthesis on a sequence of characters where you've observed an opening parenthesis. There it is a well understood problem.

For finance it's not clear to me whether there is a benefit. It may well be they could capture some sort of market regime; similar to a Markov switching process. I've never found it to be that useful, and it doesn't seem to be well-suited to identifying these sort of longer-term effects with very noisy data. There is quite a bit of work to be done on this.

What are the input and output variables in your example?

The input variables are the daily log returns of each stock in the index where we have a complete history of prices. The output is driven by the direction of the return of an equally weighted portfolio on the next day [as detailed above].

As discussed above there are various considerations such as transaction costs which need to be considered, but the idea here is simply to illustrate how high-dimensional data can translate into decision-making at some aggregate level.

The model is trained over a rolling window, which is important. If you don't retrain you will see a downward drift in performance. How frequently you retrain is an open question. It is difficult when you have knowledge of the market's history not to bias your data gathering in some way.

What was the network architecture you used?

This was a very simple architecture with three hidden layers.

Given deep learning is very "data-greedy", is there any hope that this could be applied over a longer-term horizon?

There is a problem with longer-term prediction. Generally I have found that the performance of this classifier (i.e. is the market going to move up/down or remain stationary), has an exponential decay in the accuracy with increased time horizon. This, of course, probably comes as no surprise. It's why, for example, weather prediction tends to be pretty bad a week ahead as opposed to being good in the near future.

Another aspect is if you want to predict at weekly intervals is the granularity of your data. Do you fit using weekly data if you are trying to predict one week ahead? This would push you away from deep learning as you would have too little data.

One could keep all the daily data, but still make a weekly forecast. The timeframe of the prediction is different from the granularity of the data: this is a possibility. This returns to the first point: prediction accuracy. One might want to include variables to indicate day of the week for example. There would need to more to the input feature space than just prices.

One can explore quite aggressively with deep learning a range of ideas of how to augment the input space in a way that it makes sense. There could be seasonality factors; there might be a variable which encodes the number of days before an earnings announcement or something else more qualitative in its style.

Are there any analogues in deep learning architectures with principal components (PCA) or partial least squares (PLS) which are more traditional approaches to dimensionality reduction?

We are starting to realize that a lot of techniques that are used in finance, speaking generally, such as linear regression, PCA or PLS are just shallow learners: learners with a single layer. And they have served us well in many applications.

The advantage of deep learners is that by adding more layers we increase the expressability of the model. The theory has not been worked out yet, but it does seem that deep learning is at least a hierarchal version of factor models and various other types of models that we already use.

UBS Equity Quantitative Research publications

Monographs, Keys and Q-Series

Title	Date
What is your fund's capacity?	Sep-17
Lower turnover and smart beta factors	Sep-17
Timing style rotations in China's domestic market	Jul-17
Should smart beta factors be orthogonalised?	Jul-17
How to pick stocks in the MSCI China A-Share Market	Jun-17
Active vs Passive: What is the Future of Active Management?	Jun-17
The hidden value in fund holdings	Jun-17
Can social network analysis enhance strategies following trading by corporate insiders?	Jun-17
Using Trend & Carry to time Global Bond Markets	May-17
What you need to know about Japanese equities	Apr-17
What times the bond market?	Mar-17
Passive Opportunities for Active Managers	Feb-17
Active vs Passive: How Will the World of Investing Evolve?	Jan-17
What will demographics mean for growth and stock market returns?	Jan-17
How to pick stocks in China's domestic market	Jan-17
Systematic Strategies for Single-Stock Futures	Oct-16
Irrational asset management	Oct-16
China domestic market – alpha for quantitative investors	Oct-16
Are you already timing styles successfully?	Sep-16
Do low-volatility stocks have interest-rate risk?	Sep-16
What does splitting the financials sector change?	Aug-16
Harvesting Yield from Cross-Asset Carry	Aug-16
When is the stock market likely to correct?	Aug-16
Is it easier to be a quant in small cap?	Aug-16
Follow the smart money	Jul-16
How can supply chains improve earnings visibility?	Jul-16
Where are the attractive dividend paying stocks?	Mar-16
Why does increasing volatility matter?	Feb-16
What crowded positions are bubbling up in equity markets	Feb-16
What happened to Value, and when will it return?	Jan-16
Who benefits from automation?	Nov-15
The Spectre of Equity-Bond allocation	Nov-15

Academic Research Monitor

Topic	Date
Explaining factor returns	Oct-17
Low-Risk, Low-Volatility, Low-Correlation patterns	Jul-17
Stories about Active Management	Jun-17
UBS UK Quantitative Conference 2017 Highlights	Apr-17
Recession and Tail Risk?	Mar-17
Where does Volatility Targeting Work?	Jan-17
ESG Quant Investing	Dec-16
Quality, Low-Risk and Momentum Investing	Nov-16
Combining Smart Beta Factors	Sep-16
Portfolio Construction and Overfitting	Jul-16
UBS Equity Markets Conference	May-16
European Quantitative Conference 2016 Highlights	Apr-16
Does Oil matter for Equity Markets?	Mar-16
Low Risk Investing	Feb-16
Value Investing	Dec-15
Analyst Forecasts and Measuring Distance	Nov-15
UBS Market Microstructure Conference	Oct-15
Equity Risk Premium Forecasting and Market Timing	Sep-15
Behavioural Investing Patterns	Jul-15
Quality and Size Investing	May-15
European Quantitative Conference 2015 Highlights	Apr-15
Smart Beta, Factors and Style Investing	Feb-15
Momentum-Investing	Jan-15
Investment Strategies & Textual Analysis Signals	Dec-14
Commodity Risk & Institutional Investing Habits	Nov-14
Index Membership, Investor (in)attention to News & Spurious Correlations	Sep-14
Forecasting the Equity Risk Premium	Aug-14
Implied Cost of Capital & Shorting Premium	Jun-14
Trend Following	Mar-14

PAS User Guides

PAS Macros	Feb-16	Reports	Apr-14
Quick Reference Guide	Nov-15	Risk Parity	Feb-13
Risk Parity and Composite Assets	Jan-15	Advanced Analysis	Oct-12
Introduction to the UBS Portfolio Analysis System	Jan-15	Risk Models	Nov-11
Long-Short Analysis	Jan-15	UBS Hybrid Risk Model	Dec-10
Installation	May-14	Quick Portfolio Analysis	Jul-10

R Advice

Conventional active returns can be misleading	Sep-17	R and Excel	Dec-16
Predictive modelling with caret	Jul-17	Rolling window calculations – which package to use	Sep-16
Tidy data science with the tidyverse	May-17	Getting started with random forests	Aug-16
Bayesian regressions with stan	Mar-17	Optimising in Rs	Jun-16
data.table, the best package in the world?	Mar-17	Speeding up R / Plotting correlation matrices	Jun-16

Team

UK – London

Maylan Cheung	+44-20-7568 4477
Ian Francis	+44-20-7568 1872
Josie Gerken	+44-20-7568 3560
Simon Iley	+44-20-7568 6327
Desi Ivanova	+44-20-7568-1754
David Jessop	+44-20-7567 9882
Claire Jones	+44-20-7568 1873
Manoj Kothari	+44-20-7568 1997
Simon Stoye	+44-20-7568 1876
Christine Vargas	+44-20-7568 2409

Hong Kong

Cathy Fang (Shanghai)	+86-021-3866 8891
Josh Holcroft	+852-2971 7705
Shanle Wu	+852-2971 7513

Australia– Sydney

Oliver Antrobus	+61-3-9242 6467
Luke Brown	+61-2-9324 3620
Pieter Stoltz	+61-2-9324 3779
Paul Winter	+61-2-9324 2080
Nathan Luk	+61-2-9324 2247

Required Disclosures

This report has been prepared by UBS Limited, an affiliate of UBS AG. UBS AG, its subsidiaries, branches and affiliates are referred to herein as UBS.

For information on the ways in which UBS manages conflicts and maintains independence of its research product; historical performance information; and certain additional disclosures concerning UBS research recommendations, please visit www.ubs.com/disclosures. The figures contained in performance charts refer to the past; past performance is not a reliable indicator of future results. Additional information will be made available upon request. UBS Securities Co. Limited is licensed to conduct securities investment consultancy businesses by the China Securities Regulatory Commission. UBS acts or may act as principal in the debt securities (or in related derivatives) that may be the subject of this report. This recommendation was finalized on: 19 October 2017 02:18 PM GMT. UBS has designated certain Research department members as Derivatives Research Analysts where those department members publish research principally on the analysis of the price or market for a derivative, and provide information reasonably sufficient upon which to base a decision to enter into a derivatives transaction. Where Derivatives Research Analysts co-author research reports with Equity Research Analysts or Economists, the Derivatives Research Analyst is responsible for the derivatives investment views, forecasts, and/or recommendations.

Analyst Certification: Each research analyst primarily responsible for the content of this research report, in whole or in part, certifies that with respect to each security or issuer that the analyst covered in this report: (1) all of the views expressed accurately reflect his or her personal views about those securities or issuers and were prepared in an independent manner, including with respect to UBS, and (2) no part of his or her compensation was, is, or will be, directly or indirectly, related to the specific recommendations or views expressed by that research analyst in the research report.

UBS Investment Research: Global Equity Rating Definitions

12-Month Rating	Definition	Coverage[1]	IB Services[2]
Buy	FSR is > 6% above the MRA.	45%	26%
Neutral	FSR is between -6% and 6% of the MRA.	39%	23%
Sell	FSR is > 6% below the MRA.	16%	11%
Short-Term Rating	**Definition**	**Coverage[3]**	**IB Services[4]**
Buy	Stock price expected to rise within three months from the time the rating was assigned because of a specific catalyst or event.	<1%	<1%
Sell	Stock price expected to fall within three months from the time the rating was assigned because of a specific catalyst or event.	<1%	<1%

Source: UBS. Rating allocations are as of 30 September 2017.
1:Percentage of companies under coverage globally within the 12-month rating category.
2:Percentage of companies within the 12-month rating category for which investment banking (IB) services were provided within the past 12 months.
3:Percentage of companies under coverage globally within the Short-Term rating category.
4:Percentage of companies within the Short-Term rating category for which investment banking (IB) services were provided within the past 12 months.

KEY DEFINITIONS:Forecast Stock Return (FSR) is defined as expected percentage price appreciation plus gross dividend yield over the next 12 months. **Market Return Assumption (MRA)** is defined as the one-year local market interest rate plus 5% (a proxy for, and not a forecast of, the equity risk premium). **Under Review (UR)** Stocks may be flagged as UR by the analyst, indicating that the stock's price target and/or rating are subject to possible change in the near term, usually in response to an event that may affect the investment case or valuation. **Short-Term Ratings** reflect the expected near-term (up to three months) performance of the stock and do not reflect any change in the fundamental view or investment case. **Equity Price Targets** have an investment horizon of 12 months.

EXCEPTIONS AND SPECIAL CASES:UK and European Investment Fund ratings and definitions are: Buy: Positive on factors such as structure, management, performance record, discount; **Neutral:** Neutral on factors such as structure, management, performance record, discount; **Sell:** Negative on factors such as structure, management, performance record, discount. **Core Banding Exceptions (CBE):** Exceptions to the standard +/-6% bands may be granted by the Investment Review Committee (IRC). Factors considered by the IRC include the stock's volatility and the credit spread of the respective company's debt. As a result, stocks deemed to be very high or low risk may be subject to higher or lower bands as they relate to the rating. When such exceptions apply, they will be identified in the Company Disclosures table in the relevant research piece.

Research analysts contributing to this report who are employed by any non-US affiliate of UBS Securities LLC are not registered/qualified as research analysts with FINRA. Such analysts may not be associated persons of UBS Securities LLC and therefore are not subject to the FINRA restrictions on communications with a subject company, public appearances, and trading securities held by a research analyst account. The name of each affiliate and analyst employed by that affiliate contributing to this report, if any, follows.

UBS Limited: David Jessop; Josie Gerken, PhD; Desi Ivanova; Claire Jones, CFA. **UBS Securities Australia Ltd:** Oliver Antrobus, CFA; Luke Brown, CFA; Pieter Stoltz; Paul Winter. **UBS Securities Co. Limited:** Cathy Fang, PhD. **UBS AG Hong Kong Branch:** Josh Holcroft; Shanle Wu, PhD.

Unless otherwise indicated, please refer to the Valuation and Risk sections within the body of this report. For a complete set of disclosure statements associated with the companies discussed in this report, including information on valuation and risk, please contact UBS Securities LLC, 1285 Avenue of Americas, New York, NY 10019, USA, Attention: Investment Research.

Global Disclaimer

This document has been prepared by UBS Limited, an affiliate of UBS AG. UBS AG, its subsidiaries, branches and affiliates are referred to herein as UBS.

Global Research is provided to our clients through UBS Neo, in certain instances, UBS.com and any other system, or distribution method specifically identified in one or more communications distributed through UBS Neo or UBS.com as an approved means for distributing Global Research (each a "System"). It may also be made available through third party vendors and distributed by UBS and/or third parties via e-mail or alternative electronic means. The level and types of services provided by Global Research to a client may vary depending upon various factors such as a client's individual preferences as to the frequency and manner of receiving communications, a client's risk profile and investment focus and perspective (e.g., market wide, sector specific, long-term, short-term, etc.), the size and scope of the overall client relationship with UBS and legal and regulatory constraints.

All Global Research is available on UBS Neo. Please contact your UBS sales representative if you wish to discuss your access to UBS Neo.

When you receive Global Research through a System, your access and/or use of such Global Research is subject to this Global Research Disclaimer and to the terms of use governing the applicable System.

When you receive Global Research via a third party vendor, e-mail or other electronic means, your use shall be subject to this Global Research Disclaimer and to UBS's Terms of Use/Disclaimer (http://www.ubs.com/global/en/legalinfo2/disclaimer.html). By accessing and/or using Global Research in this manner, you are indicating that you have read and agree to be bound by our Terms of Use/Disclaimer. In addition, you consent to UBS processing your personal data and using cookies in accordance with our Privacy Statement (http://www.ubs.com/global/en/legalinfo2/privacy.html) and cookie notice (http://www.ubs.com/global/en/homepage/cookies/cookie-management.html).

If you receive Global Research, whether through a System or by any other means, you agree that you shall not copy, revise, amend, create a derivative work, transfer to any third party, or in any way commercially exploit any UBS research provided via Global Research or otherwise, and that you shall not extract data from any research or estimates provided to you via Global Research or otherwise, without the prior written consent of UBS.

This document is for distribution only as may be permitted by law. It is not directed to, or intended for distribution to or use by, any person or entity who is a citizen or resident of or located in any locality, state, country or other jurisdiction where such distribution, publication, availability or use would be contrary to law or regulation or would subject UBS to any registration or licensing requirement within such jurisdiction. It is published solely for information purposes; it is not an advertisement nor is it a solicitation or an offer to buy or sell any financial instruments or to participate in any particular trading strategy. No representation or warranty, either expressed or implied, is provided in relation to the accuracy, completeness or reliability of the information contained in any materials to which this document relates (the "Information"), except with respect to Information concerning UBS. The Information is not intended to be a complete statement or summary of the securities, markets or developments referred to in the document. UBS does not undertake to update or keep current the Information. Any opinions expressed in this document may change without notice and may differ or be contrary to opinions expressed by other business areas or groups, personnel or other representative of UBS. Any statements contained in this report attributed to a third party represent UBS's interpretation of the data, information and/or opinions provided by that third party either publicly or through a subscription service, and such use and interpretation have not been reviewed by the third party.

This document is a general communication and is educational in nature. Nothing in this document constitutes a representation that any investment strategy or recommendation is suitable or appropriate to an investor's individual circumstances or otherwise constitutes a personal recommendation. By providing this document, none of UBS or its representatives has any responsibility or authority to provide or have provided investment advice in a fiduciary capacity or otherwise. Investments involve risks, and investors should exercise prudence and their own judgement in making their investment decisions. None of UBS or its representatives is suggesting that the recipient or any other person take a specific course of action or any action at all. By receiving this document, the recipient acknowledges and agrees with the intended purpose described above and further disclaims any expectation or belief that the information constitutes investment advice to the recipient or otherwise purports to meet the investment objectives of the recipient. The financial instruments described in the document may not be eligible for sale in all jurisdictions or to certain categories of investors. Options, derivative products and futures are not suitable for all investors, and trading in these instruments is considered risky. Mortgage and asset-backed securities may involve a high degree of risk and may be highly volatile in response to fluctuations in interest rates or other market conditions. Foreign currency rates of exchange may adversely affect the value, price or income of any security or related instrument referred to in the document. For investment advice, trade execution or other enquiries, clients should contact their local sales representative.

The value of any investment or income may go down as well as up, and investors may not get back the full (or any) amount invested. Past performance is not necessarily a guide to future performance. Neither UBS nor any of its directors, employees or agents accepts any liability for any loss (including investment loss) or damage arising out of the use of all or any of the Information.

Any prices stated in this document are for information purposes only and do not represent valuations for individual securities or other financial instruments. There is no representation that any transaction can or could have been effected at those prices, and any prices do not necessarily reflect UBS's internal books and records or theoretical model-based valuations and may be based on certain assumptions. Different assumptions by UBS or any other source may yield substantially different results.

This document and the Information are produced by UBS as part of its research function and are provided to you solely for general background information. UBS has no regard to the specific investment objectives, financial situation or particular needs of any specific recipient. In no circumstances may this document or any of the Information be used for any of the following purposes:

(i) valuation or accounting purposes;

(ii) to determine the amounts due or payable, the price or the value of any financial instrument or financial contract; or

(iii) to measure the performance of any financial instrument.

By receiving this document and the Information you will be deemed to represent and warrant to UBS that you will not use this document or any of the Information for any of the above purposes or otherwise rely upon this document or any of the Information. Prior to making any investment or financial decisions, any recipient of this document or the information should seek individualized advice from his or her personal financial, legal, tax and other professional advisors that takes into account all the particular facts and circumstances of his or her investment objectives.

UBS has policies and procedures, which include, without limitation, independence policies and permanent information barriers, that are intended, and upon which UBS relies, to manage potential conflicts of interest and control the flow of information within divisions of UBS and among its subsidiaries, branches and affiliates. For further information on the ways in which UBS manages conflicts and maintains independence of its research products, historical performance information and certain additional disclosures concerning UBS research recommendations, please visit www.ubs.com/disclosures.

Research will initiate, update and cease coverage solely at the discretion of UBS Research Management, which will also have sole discretion on the timing and frequency of any published research product. The analysis contained in this document is based on numerous assumptions. All material information in relation to published research reports, such as valuation methodology, risk statements, underlying assumptions (including sensitivity analysis of those assumptions), ratings history etc. as required by the Market Abuse Regulation, can be found on NEO. Different assumptions could result in materially different results.

The analyst(s) responsible for the preparation of this document may interact with trading desk personnel, sales personnel and other parties for the purpose of gathering, applying and interpreting market information. UBS relies on information barriers to control the flow of information contained in one or more areas within UBS into other areas, units, groups or affiliates of UBS. The compensation of the analyst who prepared this document is determined exclusively by research management and senior management (not including investment banking). Analyst compensation is not based on investment banking revenues; however, compensation may relate to the revenues of UBS and/or its divisions as a whole, of which investment banking, sales and trading are a part, and UBS's subsidiaries, branches and affiliates as a whole.

For financial instruments admitted to trading on an EU regulated market: UBS AG, its affiliates or subsidiaries (excluding UBS Securities LLC) acts as a market maker or liquidity provider (in accordance with the interpretation of these terms in the UK) in the financial instruments of the issuer save that where the activity of liquidity provider is carried out in accordance with the definition given to it by the laws and regulations of any other EU jurisdictions, such information is separately disclosed in this document. For financial instruments admitted to trading on a non-EU regulated market: UBS may act as a market maker save that where this activity is carried out in the US in accordance with the definition given to it by the relevant laws and regulations, such activity will be specifically disclosed in this document. UBS may have issued a warrant the value of which is based on one or more of the financial instruments referred to in the document. UBS and its affiliates and employees may have long or short positions, trade as principal and buy and sell in instruments or derivatives identified herein; such transactions or positions may be inconsistent with the opinions expressed in this document.

United Kingdom and the rest of Europe: Except as otherwise specified herein, this material is distributed by UBS Limited to persons who are eligible counterparties or professional clients. UBS Limited is authorised by the Prudential Regulation Authority and regulated by the Financial Conduct Authority and the Prudential Regulation Authority. **France:** Prepared by UBS Limited and distributed by UBS Limited and UBS Securities France S.A. UBS Securities France S.A. is regulated by the ACPR (Autorité de Contrôle Prudentiel et de Résolution) and the Autorité des Marchés Financiers (AMF). Where an analyst of UBS Securities France S.A. has contributed to this document, the document is also deemed to have been prepared by UBS Securities France S.A. **Germany:** Prepared by UBS Limited and distributed by UBS Limited and UBS Europe SE. UBS Europe SE is regulated by the Bundesanstalt fur Finanzdienstleistungsaufsicht (BaFin). **Spain:** Prepared by UBS Limited and distributed by UBS Limited and UBS Securities España SV, SA. UBS Securities España SV, SA is regulated by the Comisión Nacional del Mercado de Valores (CNMV). **Turkey:** Distributed by UBS Limited. No information in this document is provided for the purpose of offering, marketing and sale by any means of any capital market instruments and services in the Republic of Turkey. Therefore, this document may not be considered as an offer made or to be made to residents of the Republic of Turkey. UBS Limited is not licensed by the Turkish Capital Market Board under the provisions of the Capital Market Law (Law No. 6362). Accordingly, neither this document nor any other offering material related to the instruments/services may be utilized in connection with providing any capital market services to persons within the Republic of Turkey without the prior approval of the Capital Market Board. However, according to article 15 (d) (ii) of the Decree No. 32, there is no restriction on the purchase or sale of the securities abroad by residents of the Republic of Turkey. **Poland:** Distributed by UBS Limited (spolka z ograniczona odpowiedzialnoscia) Oddzial w Polsce regulated by the Polish Financial Supervision Authority. Where an analyst of UBS Limited (spolka z ograniczona odpowiedzialnoscia) Oddzial w Polsce has contributed to this document, the document is also deemed to have been prepared by UBS Limited (spolka z ograniczona odpowiedzialnoscia) Oddzial w Polsce. **Russia:** Prepared and distributed by UBS Bank (OOO). **Switzerland:** Distributed by UBS AG to persons who are institutional investors only. UBS AG is regulated by the Swiss Financial Market Supervisory Authority (FINMA). **Italy:** Prepared by UBS Limited and distributed by UBS Limited and UBS Limited, Italy Branch. Where an analyst of UBS Limited, Italy Branch has contributed to this document, the document is also deemed to have been prepared by UBS Limited, Italy Branch. **South Africa:** Distributed by UBS South Africa (Pty) Limited (Registration No. 1995/011140/07), an authorised user of the JSE and an authorised Financial Services Provider (FSP 7328). **Israel:** This material is distributed by UBS Limited. UBS Limited is authorised by the Prudential Regulation Authority and regulated by the Financial Conduct Authority and the Prudential Regulation Authority. UBS Securities Israel Ltd is a licensed Investment Marketer that is supervised by the Israel Securities Authority (ISA). UBS Limited and its affiliates incorporated outside Israel are not licensed under the Israeli Advisory Law. UBS Limited is not covered by insurance as required from a licensee under the Israeli Advisory Law. UBS may engage among others in issuance of Financial Assets or in distribution of Financial Assets of other issuers for fees or other benefits. UBS Limited and its affiliates may prefer various Financial Assets to which they have or may have Affiliation (as such term is defined under the Israeli Advisory Law). Nothing in this Material should be considered as investment advice under the Israeli Advisory Law. This Material is being issued only to and/or is directed only at persons who are Eligible Clients within the meaning of the Israeli Advisory Law, and this material must not be relied on or acted upon by any other persons. **Saudi Arabia:** This document has been issued by UBS AG (and/or any of its subsidiaries, branches or affiliates), a public company limited by shares, incorporated in Switzerland with its registered offices at Aeschenvorstadt 1, CH-4051 Basel and Bahnhofstrasse 45, CH-8001 Zurich. This publication has been approved by UBS Saudi Arabia (a subsidiary of UBS AG), a Saudi closed joint stock company incorporated in the Kingdom of Saudi Arabia under commercial register number 1010257812 having its registered office at Tatweer Towers, P.O. Box 75724, Riyadh 11588, Kingdom of Saudi Arabia. UBS Saudi Arabia is authorized and regulated by the Capital Market Authority to conduct securities business under license number 08113-37. **UAE / Dubai:** The information distributed by UBS AG Dubai Branch is only intended for Professional Clients and/or Market Counterparties, as classified under the DFSA rulebook. No other person should act upon this material/communication. The information is not for further distribution within the United Arab Emirates. UBS AG Dubai Branch is regulated by the DFSA in the DIFC. UBS is not licensed to provide banking services in the UAE by the Central Bank of the UAE, nor is it licensed by the UAE Securities and Commodities Authority. **United States:** Distributed to US persons by either UBS Securities LLC or by UBS Financial Services Inc., subsidiaries of UBS AG; or by a group, subsidiary or affiliate of UBS AG that is not registered as a US broker-dealer (a **'non-US affiliate'**) to major US institutional investors only. UBS Securities LLC or UBS Financial Services Inc. accepts responsibility for the content of a document prepared by another non-US affiliate when distributed to US persons by UBS Securities LLC or UBS Financial Services Inc. All transactions by a US person in the securities mentioned in this document must be effected through UBS Securities LLC or UBS Financial Services Inc., and not through a non-US affiliate. UBS Securities LLC is not acting as a municipal advisor to any municipal entity or obligated person within the meaning of Section 15B of the Securities Exchange Act (the "Municipal Advisor Rule"), and the opinions or views contained herein are not intended to be, and do not constitute, advice within the meaning of the Municipal Advisor Rule. **Canada:** Distributed by UBS Securities Canada Inc., a registered investment dealer in Canada and a Member-Canadian Investor Protection Fund, or by another affiliate of UBS AG that is registered to conduct business in Canada or is otherwise exempt from registration. **Mexico:** This report has been distributed and prepared by UBS Casa de Bolsa, S.A. de C.V., UBS Grupo Financiero, an entity that is part of UBS Grupo Financiero, S.A. de C.V. and is a subsidiary of UBS AG. This document is intended for distribution to institutional or sophisticated investors only. Research reports only reflect the views of the analysts responsible for the reports. Analysts do not receive any compensation from persons or entities different from UBS Casa de Bolsa, S.A. de C.V., UBS Grupo Financiero, or different from entities belonging to the same financial group or business group of such. For Spanish translations of applicable disclosures, please go to www.ubs.com/disclosures. **Brazil:** Except as otherwise specified herein, this material is prepared by UBS Brasil CCTVM S.A. to persons who are eligible investors residing in Brazil, which are considered to be: (i) financial institutions, (ii) insurance firms and investment capital companies, (iii) supplementary pension entities, (iv) entities that hold financial investments higher than R$300,000.00 and that confirm the status of qualified investors in written, (v) investment funds, (vi) securities portfolio managers and securities consultants duly authorized by Comissão de Valores Mobiliários (CVM), regarding their own investments, and (vii) social security systems created by the Federal Government, States, and Municipalities. **Hong Kong:** Distributed by UBS Securities Asia Limited and/or UBS AG, Hong Kong Branch. Please contact local licensed/registered representatives of UBS Securities Asia Limited and/or UBS AG, Hong Kong Branch in respect of any matters arising from, or in connection with, the analysis or document. **Singapore:** Distributed by UBS Securities Pte. Ltd. [MCI (P) 008/09/2017 and Co. Reg. No.: 198500648C] or UBS AG, Singapore Branch. Please contact UBS Securities Pte. Ltd., an exempt financial adviser under the Singapore Financial Advisers Act (Cap. 110); or UBS AG, Singapore Branch, an exempt financial adviser under the Singapore Financial Advisers Act (Cap. 110) and a wholesale bank licensed under the Singapore Banking Act (Cap. 19) regulated by the Monetary Authority of Singapore, in respect of any matters arising from, or in connection with, the analysis or document. The recipients of this document represent and warrant that they are accredited and institutional investors as defined in the Securities and Futures Act (Cap. 289). **Japan:** Distributed by UBS Securities Japan Co., Ltd. to professional investors (except as otherwise permitted). Where this document has been prepared by UBS Securities Japan Co., Ltd., UBS Securities Japan Co., Ltd. is the author, publisher and distributor of the document. Distributed by UBS AG, Tokyo Branch to Professional Investors (except as otherwise permitted) in relation to foreign exchange and other banking businesses when relevant. **Australia:** Clients of UBS AG: Distributed by UBS AG (ABN 47 088 129 613 and holder of Australian Financial Services License No. 231087). Clients of UBS Securities Australia Ltd: Distributed by UBS Securities Australia Ltd (ABN 62 008 586 481 and holder of Australian Financial Services License No. 231098). This Document contains general information and/or general advice only and does not constitute personal financial product advice. As such, the Information in this document has been prepared without taking into account any investor's objectives, financial situation or needs, and investors should, before acting on the Information, consider the appropriateness of the Information, having regard to their objectives, financial situation and needs. If the Information contained in this document relates to the acquisition, or potential acquisition of a particular financial product by a 'Retail' client as defined by section 761G of the Corporations Act 2001 where a Product Disclosure Statement would be required, the retail client should obtain and consider the Product Disclosure Statement relating to the product before making any decision about whether to acquire the product. The UBS Securities Australia Limited Financial Services Guide is available at: www.ubs.com/ecs-research-fsg. **New Zealand:** Distributed by UBS New Zealand Ltd. UBS New Zealand Ltd is not a registered bank in New Zealand. You are being provided with this UBS publication or material because you have indicated to UBS that you are a "wholesale client" within the meaning of section 5C of the Financial Advisers Act 2008 of New Zealand (Permitted Client). This publication or material is not intended for clients who are not Permitted Clients (non-permitted Clients). If you are a non-permitted Client you must not rely on this publication or material. If despite this warning you nevertheless rely on this publication or material, you hereby (i) acknowledge that you may not rely on the content of this publication or material and that any recommendations or opinions in such this publication or material are not made or provided to you, and (ii) to the maximum extent permitted by law (a) indemnify UBS and its associates or related entities (and their respective Directors, officers, agents and Advisors) (each a 'Relevant Person') for any loss, damage, liability or claim any of them may incur or suffer as a result of, or in connection with, your unauthorised reliance on this publication or material and (b) waive any rights or remedies you may have against any Relevant Person for (or in respect of) any loss, damage, liability or claim you may incur or suffer as a result of, or in connection with, your unauthorised reliance on this publication or material. **Korea:** Distributed in Korea by UBS Securities Pte. Ltd., Seoul Branch. This document may have been edited or contributed to from time to time by affiliates of UBS Securities Pte. Ltd., Seoul Branch. This material is intended for professional/institutional clients only and not for distribution to any retail clients. **Malaysia:** This material is authorized to be distributed in Malaysia by UBS Securities Malaysia Sdn. Bhd (Capital Markets Services License No.: CMSL/A0063/2007). This material is intended for professional/institutional clients only and not for distribution to any retail clients. **India:** Distributed by UBS Securities India Private Ltd. (Corporate Identity Number U67120MH1996PTC097299) 2/F, 2 North Avenue, Maker Maxity, Bandra Kurla Complex, Bandra (East), Mumbai (India) 400051. Phone: +912261556000. It provides brokerage services bearing SEBI Registration Numbers: NSE (Capital Market Segment): INB230951431, NSE (F&O Segment) INF230951431, NSE (Currency Derivatives Segment) INE230951431, BSE (Capital Market Segment) INB010951437; merchant banking services bearing SEBI Registration Number: INM000010809 and Research Analyst services bearing SEBI Registration Number: INH000001204. UBS AG, its affiliates or subsidiaries may have debt holdings or positions in the subject Indian company/companies. Within the past 12 months, UBS AG, its affiliates or subsidiaries may have received compensation for non-investment banking securities-related services and/or non-securities services from the subject Indian company/companies. The subject company/companies may have been a client/clients of UBS AG, its affiliates or subsidiaries during the 12 months preceding the date of distribution of the research report with respect to investment banking and/or non-investment

banking securities-related services and/or non-securities services. With regard to information on associates, please refer to the Annual Report at: http://www.ubs.com/global/en/about_ubs/investor_relations/annualreporting.html

The disclosures contained in research documents produced by UBS Limited shall be governed by and construed in accordance with English law.