Our presentation is the fruit of our an ongoing research.

Topic
The implementation of a novel approach of Deep Reinforcement Learning for Market Prediction using of Big Data Analytics as a tool to enrich our financial dataset.

Key Takeaway
The audiance will be exposed to a novel architecture that uses Deep Learning as a tool for Big Data Analytics to populate a rich datset, then to be fed to a Deep Reinforcement Learning model for Trading.

# DEEP REINFORCEMENT LEARNING USING BIG DATA FOR STOCK MARKET PREDICTION

By Mouafek Ayadi

Data Scientist at ODDO BHF

July 2021

# Table of Contents

# ALGORITHMIC TRADING

## PART 1
INTRODUCTION

# ALGORITHMIC TRADING

"Algorithmic trading is a type of quant trading that uses pre-specified machine executable instructions to determine the size and timing of trades based on a quantitative model of an asset's price behavior. Over 70 percent of US trading volume is algorithmic. Most of this volume is high-frequency trading."[1]

# ALGORITHMIC TRADING

## PART 2
MACHINE LEARNING

# ALGORITHMIC TRADING
## WITH MACHINE LEARNING

"Machine learning (ML) involves algorithms that learn rules or patterns from data to achieve a goal such as minimizing a prediction error. ML algorithms can extract information from data to support or automate key investment activities.

These activities include observing the market and analyzing data to form expectations about the future and decide on placing buy or sell orders, as well as managing the resulting portfolio to produce attractive returns relative to the risk."[2]

# BIG DATA IN ALGO-TRADING

## PART 1
### BIG DATA OVERVIEW

# BIG DATA
## OVERVIEW

“Financial trading data sets reflect the myriad of decisions taken by market participants. According to Herbert Simon, actors begin their decision making processes by attempting to gather information. In today’s world, information gathering often consists of searching online sources. Recently, the search engine Google has begun to provide access to aggregated information on the volume of queries for different search terms and how these volumes change over time, via the publicly available service Google Trends.”[3]

# BIG DATA IN ALGO-TRADING

## PART 2
### WIDE & DEEP LEARNING

# BIG DATA FOR DEEP MODELS

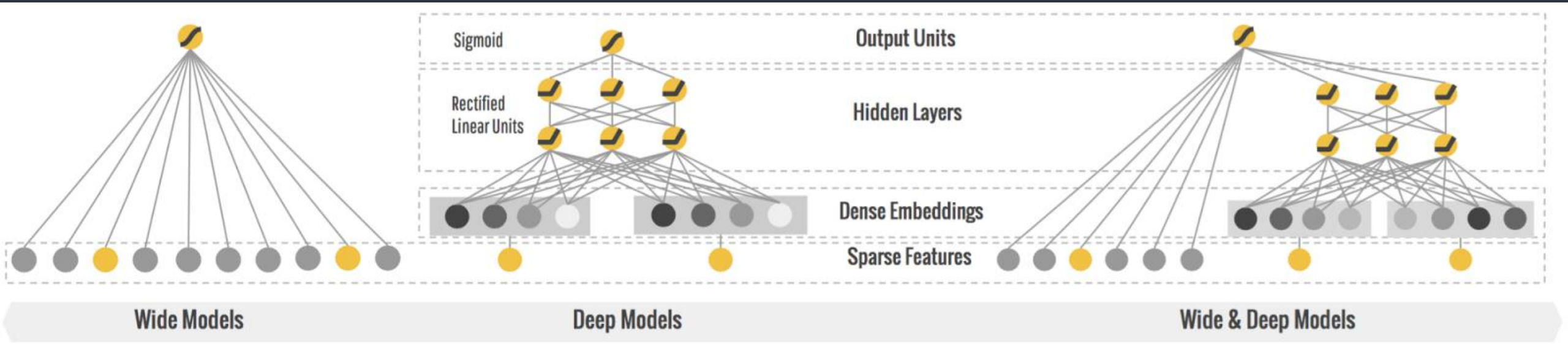
Since it’s based on Google Architecture we provide you with our findings in their literature and practice that supports our research.

“An embedding is a translation of a high-dimensional vector into a low-dimensional space. Ideally, an embedding captures some of the semantics of the input by placing semantically similar inputs close together in the embedding space.” [4]

It’s our challenge to take advantage of what’s usually used in Language modeling in our Time-Series based challenge specifically in Trading.

“This is the first work employing gradient boosting of deep models with embedding in the context of time-series forecasting.” [5]

The previous statement proves that we can use embedding in modeling Time Series, and it’s possible to create dense embeddings with Big Data.



# BIG DATA FOR WIDE MODELS

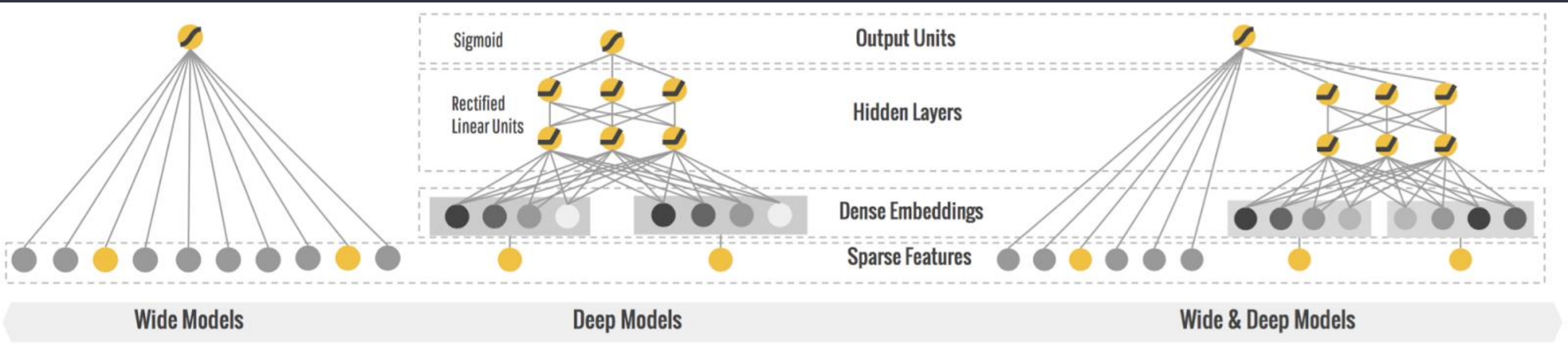
The idea is to have more quantifiable features aggregated from different resources like Google Trends, Twitter and news Articles.

We have scrapped the Google Trends data to have a better understanding of how well/bad companies are doing on the Stock Exchanges.

”The more weekly searches a company has the more it’s traded on the New York Stock Exchange.“ [6]

We have adjusted our findings using Sentiment Analysis on News Article and Tweets that we scrap on daily basis so it’s stays relevant.

“Algorithms scrape the language millions of people use on Twitter and in Google searches, determining whether people are thinking positively or negatively about a company or product.” [6]



# BIG DATA FOR WIDE & DEEP LEARNING

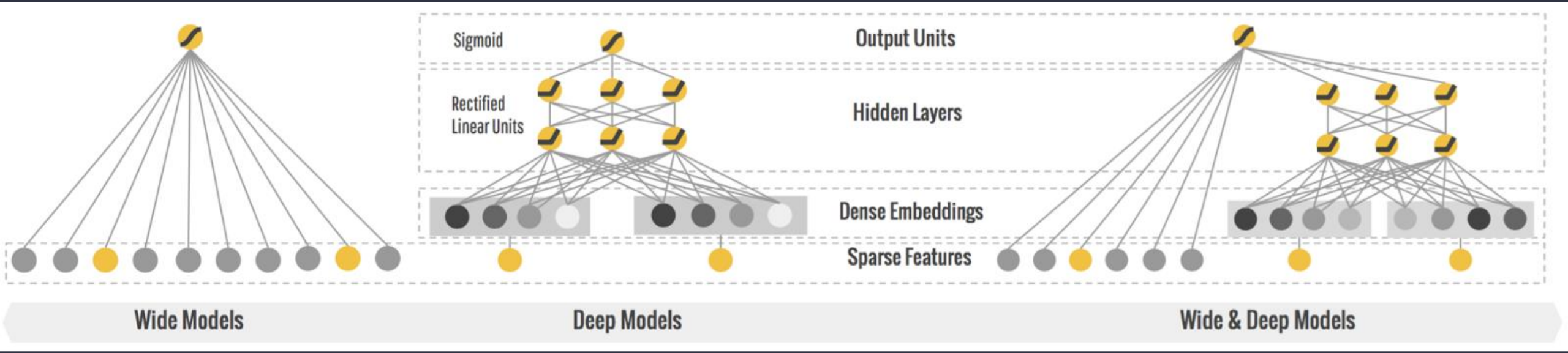
The following is the Google Architecture used in many of their products especially in Recommendation Systems and we have adapt it for Trading.

“Can we teach computers to learn like humans do, by combining the power of memorization and generalization?” [4]

Wide & Deep Learning developed at Google Research was the answer to the previous question.

“It's not an easy question to answer, but by jointly training a wide linear model (for memorization) alongside a deep neural network (for generalization), one can combine the strengths of both to bring us one step closer.” [4]

“It's useful for generic large-scale regression and classification problems with sparse inputs (categorical features with a large number of possible feature values), such as recommender systems, search, and ranking problems.” [4]



# BIG DATA IN ALGO-TRADING

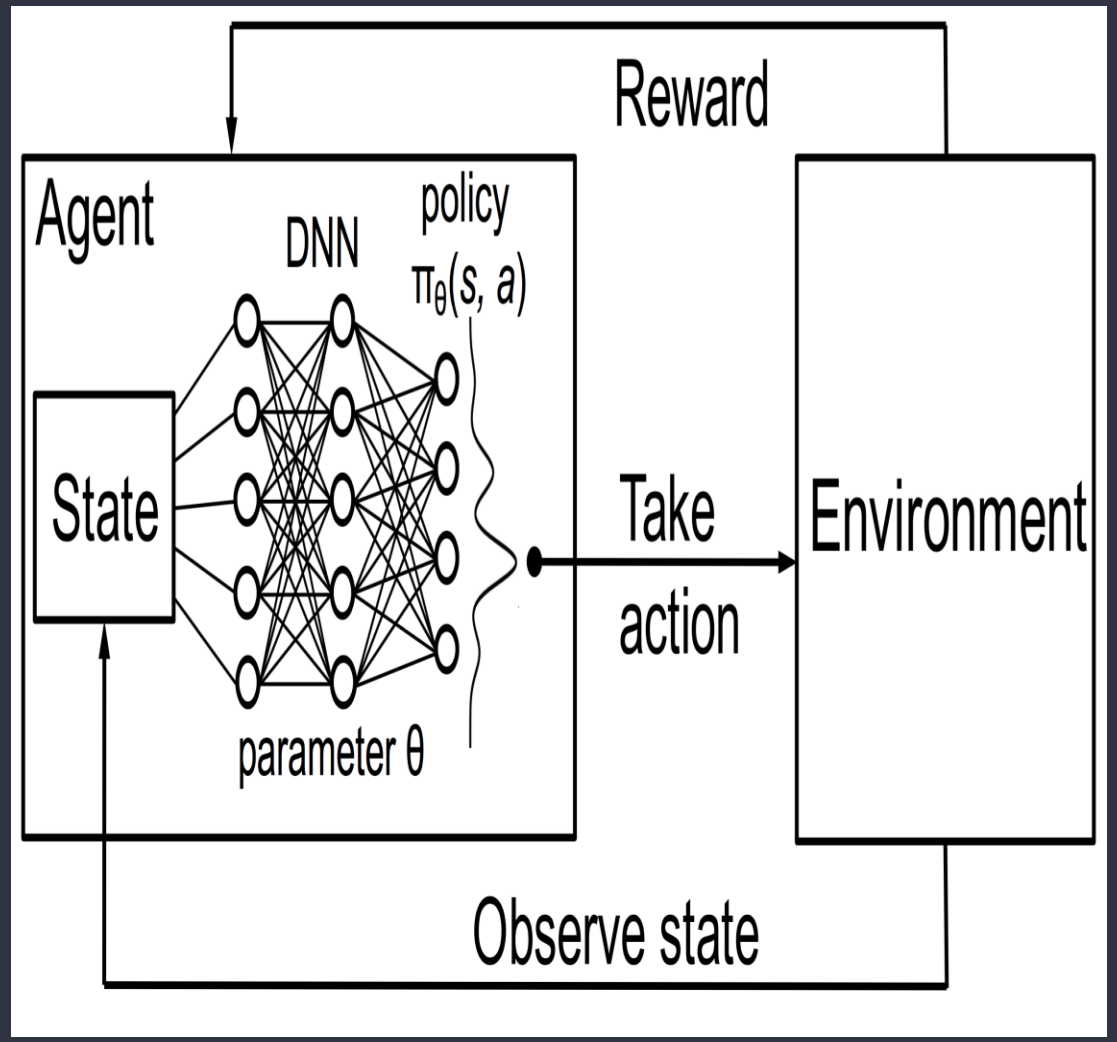
## PART 3
### DEEP REINFORCEMENT LEARNING

## DEEP REINFORCEMENT LEARNING

### I n t r o

“Deep reinforcement learning is the combination of reinforcement learning (RL) and deep learning. This field of research has been able to solve a wide range of complex decision-making tasks that were previously out of reach for a machine.

Thus, deep RL opens up many new applications in domains such as healthcare, robotics, smart grids, finance, and many more.” [5]



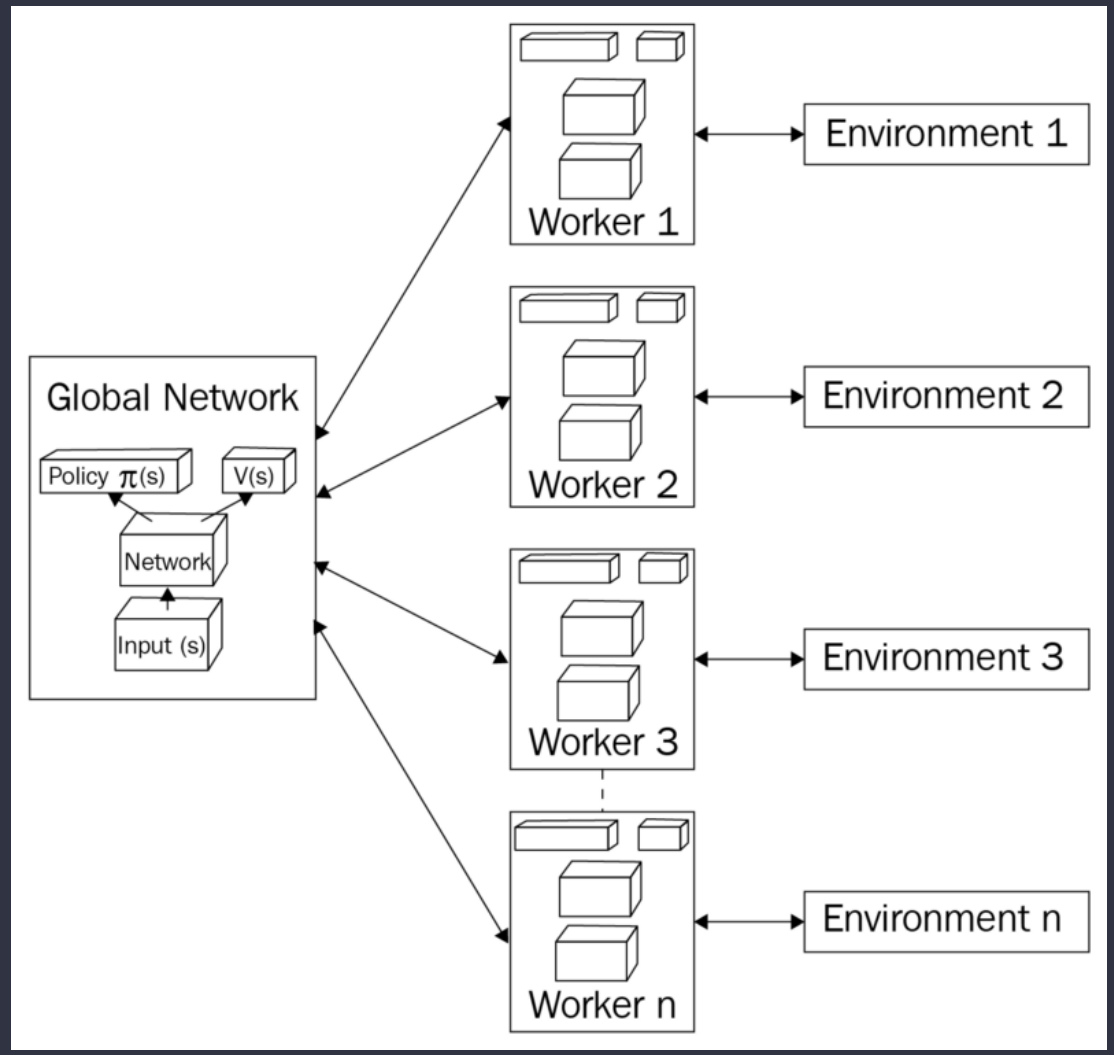
## DEEP REINFORCEMENT LEARNING

### A 3 C

The Asynchronous Advantage Actor Critic (A3C) was developed by Google's DeepMind.

A3C consists of multiple independent agents(networks) with their own weights, who interact with a different copy of the environment in parallel.

“We now describe our variants of one-step Q-learning, one-step Sarsa, n-step Q-learning and advantage actor-critic”. [6]



# OUR PROPOSED MODEL ARCHITECTURE

## PART 1
### THE EXPLANATION

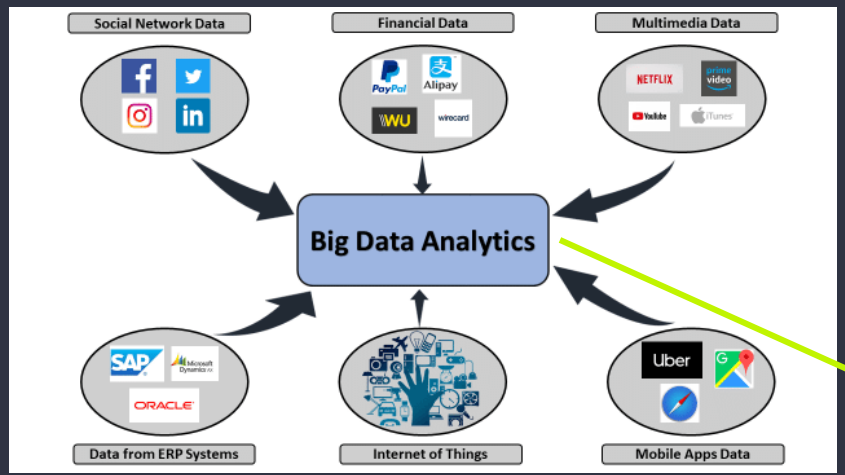
# DEEP REINFORCEMENT LEARNING FOR TRADING

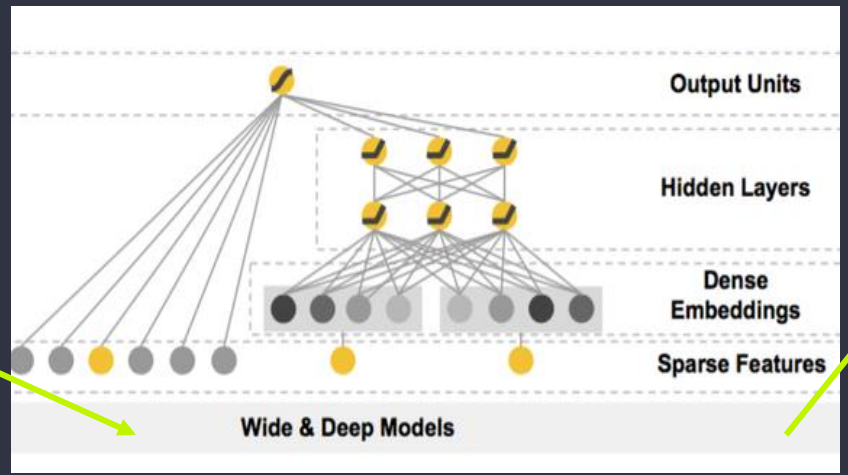
## A3C WITH WIDE & DEEP NETWORKS

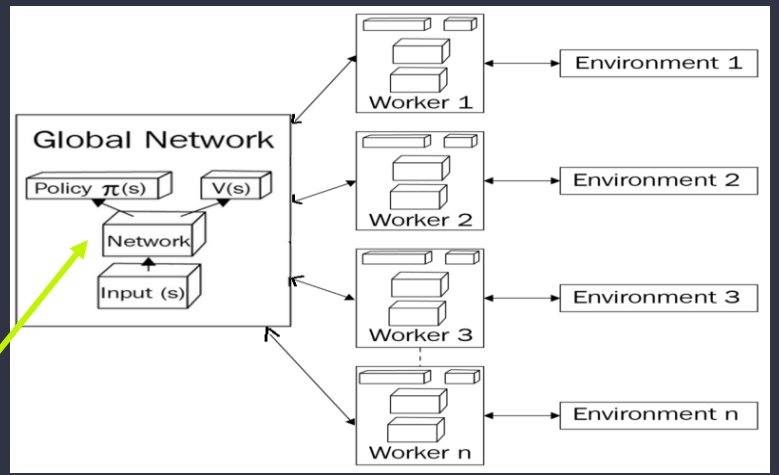
In a single financial market we can implement A3C that consists of multiple independent agents(networks) with their own weights that can operate in different sectors, who interact with a different copy of the environment in parallel used to simulate different sector scenarios.

Our idea is to make the Model-Free A3C bound by the market previous behavior and current trends, in order to predict a week ahead stock values. This procedure is accurate since it's based on Big Data endogenous and exogenous variables and factors.

In order to make the Artificial Neural Network more compliant with the Financial markets and their respective sectors where the agents operates in, we define a Wide & Deep Models, where wide linear model is used for memorization (historical market/sector pattern), alongside a deep neural network used for generalization (detect new market/sector pattern).



# OUR PROPOSED MODEL ARCHITECTURE

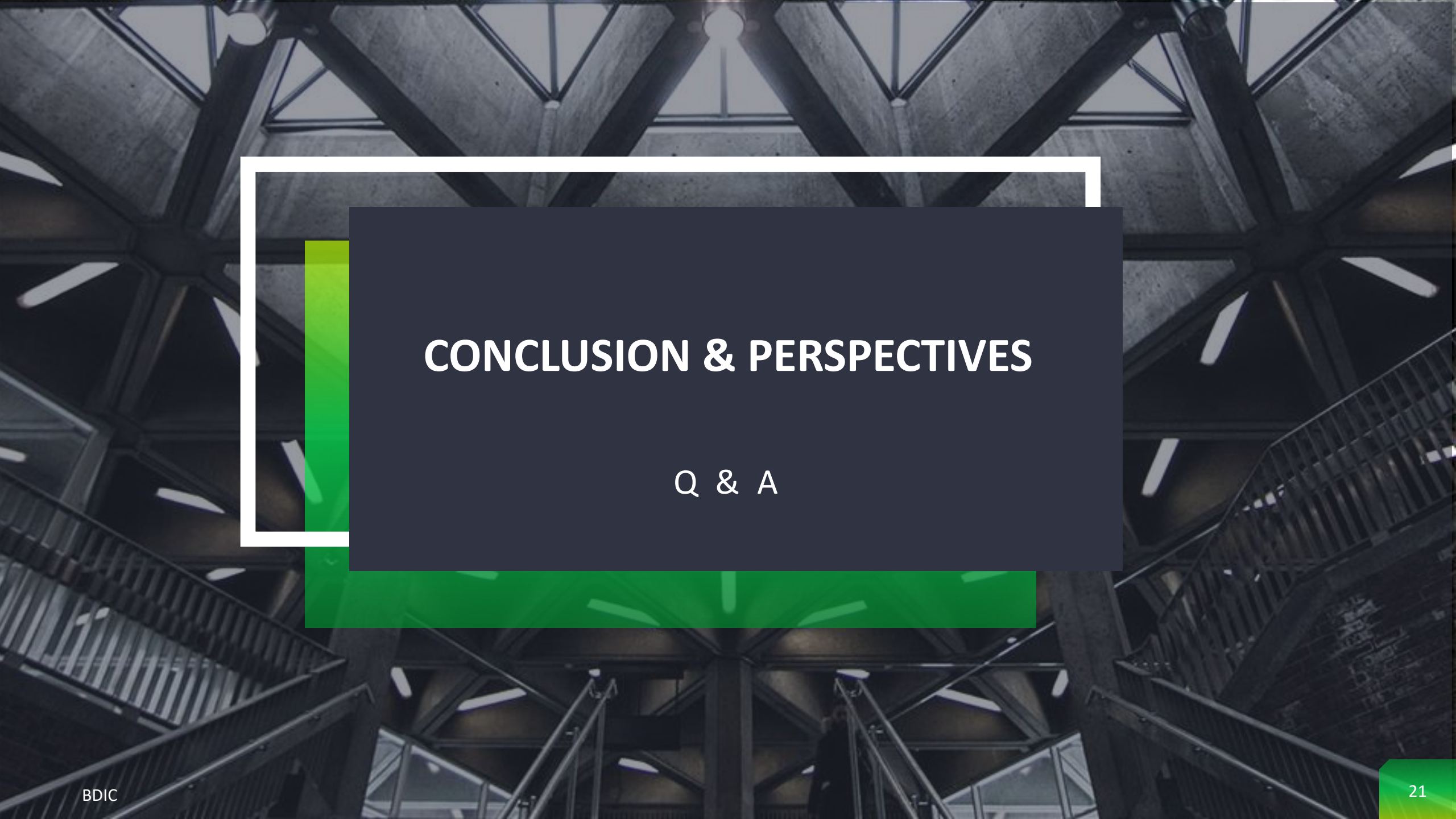
## PART 2
### THE REFERENCES

## Personal Work (ongoing)

https://github.com/MWFK/Market_Prediction-Building_Trading_Strategies

## References


[1] https://www.yuque.com/xhades/hs3xtz/ibg3cw
[2] https://ml4trading.io/chapter/0
[3] http://wrap.warwick.ac.uk/54797/14/WRAP_Preis_srep01684.pdf
[4] https://cloud.google.com/architecture/overview-extracting-and-serving-feature-embeddings-for-machine-learning
[5] https://arxiv.org/pdf/2104.04781.pdf
[6] https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2010.0284
https://www.nature.com/articles/srep01684
[7] https://arxiv.org/pdf/1606.07792.pdf
[8] https://arxiv.org/abs/1811.12560
[9] https://arxiv.org/pdf/1602.01783.pdf



# CONCLUSION & PERSPECTIVES

Q & A