What use of AI in finance?

ISCF Final Presentation

Supervisor: Prof. Dritan NACE

Denis DEMKO **Thomas DEROO** **Doris FEJZA**
Elona KARAJ **Hussein LEZZAIK** **Estia MALIQARI**
Yijue XIE

Table of Contents

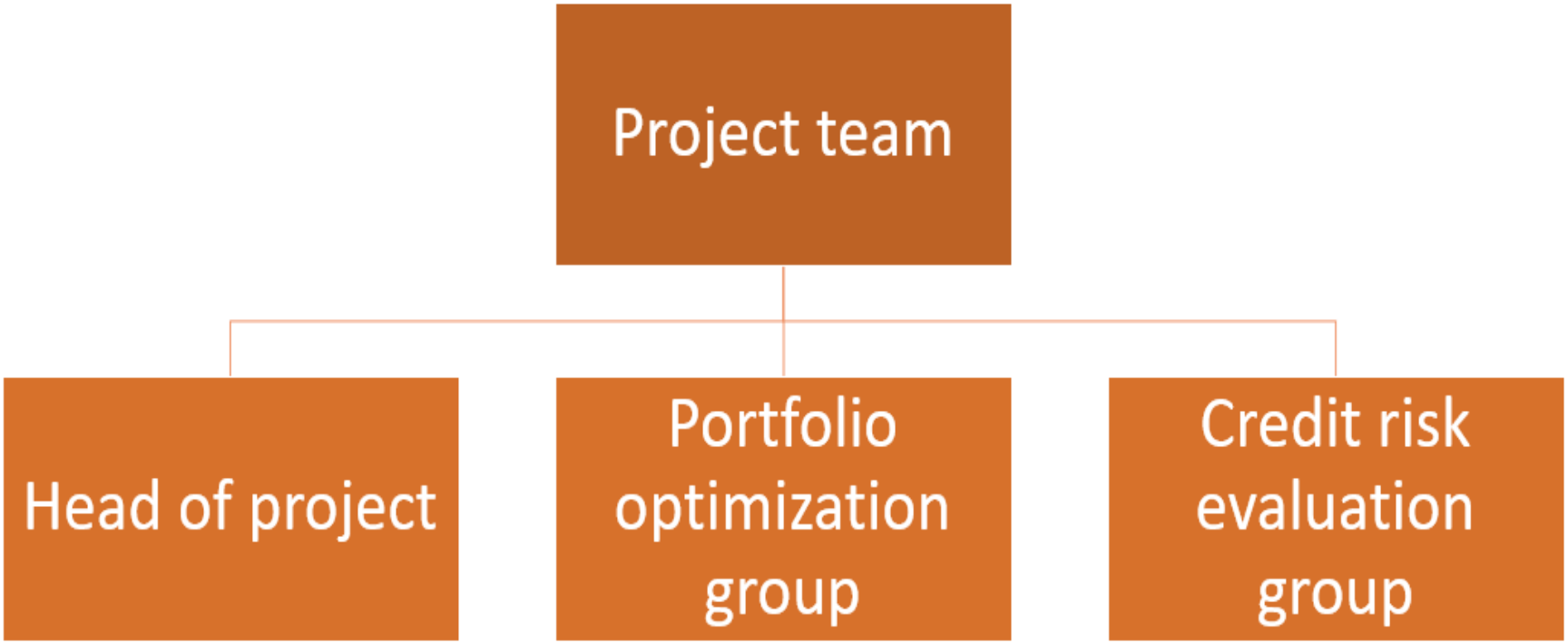
Objectives

PROVIDE A STUDY ON THE VARIOUS APPLICATIONS OF MACHINE LEARNING IN FINANCE

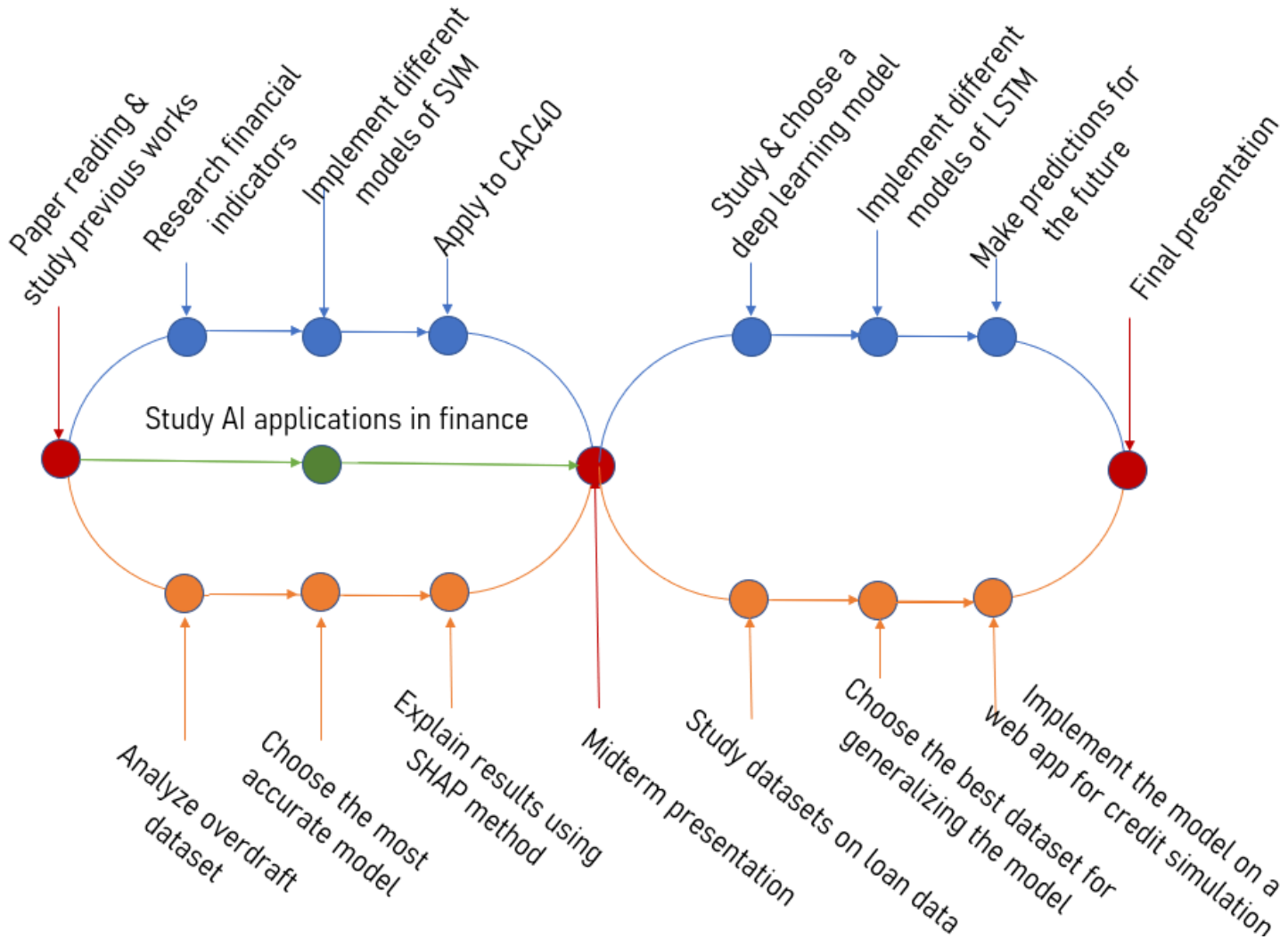
STUDY THE PORTFOLIO MANAGEMENT PROBLEM AND THE APPLICATION OF THE SVM & DEEP LEARNING METHODS

STUDY THE TOOL DEVELOPED FOR THE APPLICATION OF CREDIT RISK ASSESSMENT

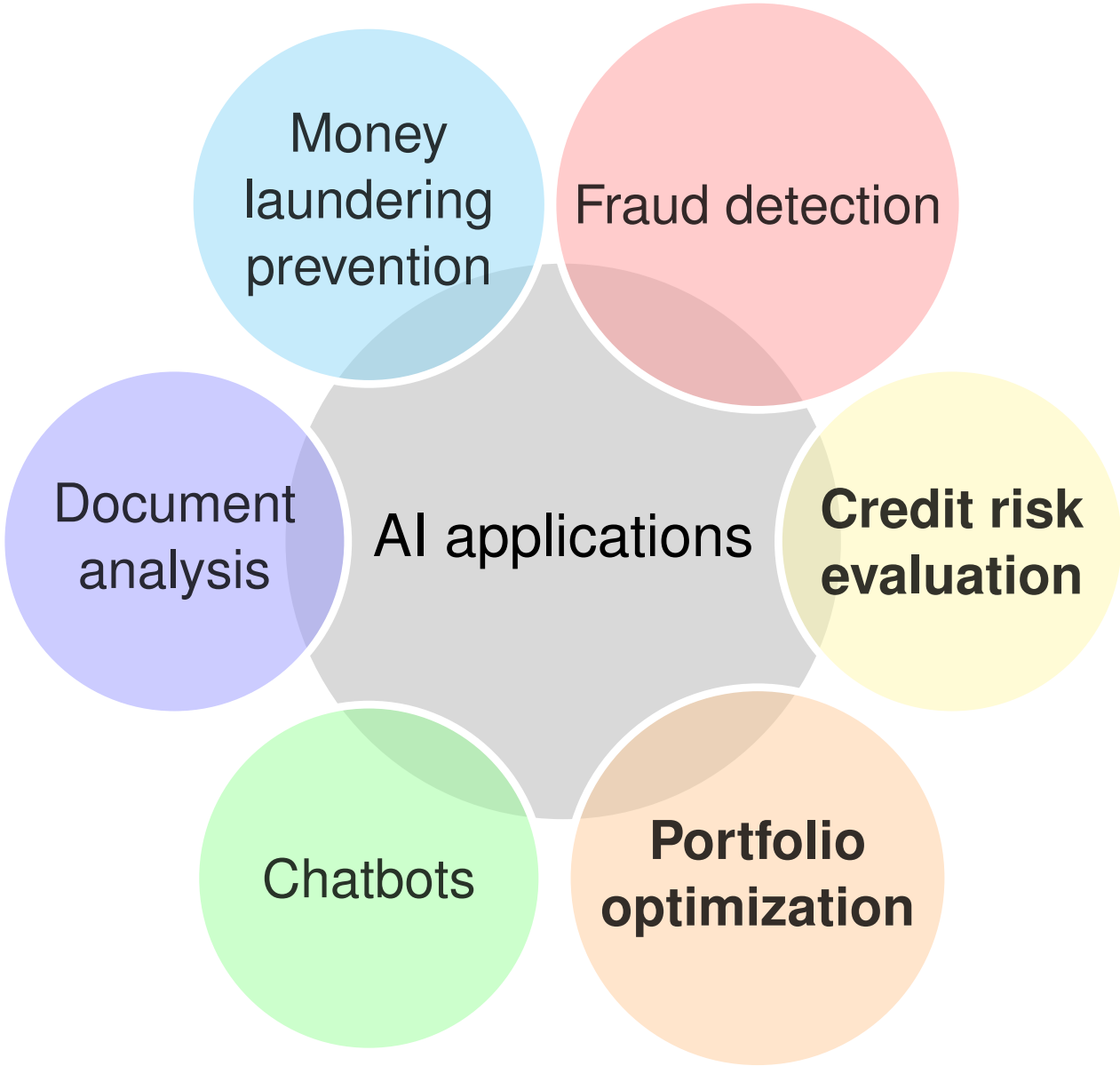
Team organization



Main milestones



AI applications in finance



Portfolio Management

Context

Objective

Provide a tool for investors to support decisions on investing in a specific stock or not, by predicting stock market behaviour.

- Implement machine learning algorithm SVM
- Implement Deep Learning algorithm LSTM
- Make 7-day predictions for market trend with LSTM

CAC40 dataset

- Source: Yahoo! Finance, FCHI ticker
- Contains information from 25/02/1990 up to the current date
- Essentially a time series with 1605 samples, 6 features

Part 1: Support Vector Machine

Predicting stock market changes

2 main theories : Fundamental analysis vs Technical Analysis

Support Vector Machines(SVM)

SVM is the machine learning algorithm used in classification problems

Our contribution

- Reading and understanding the financial and technical aspects of the project
- Adapted pre-processing steps to conform to our dataset
- Tuned learning parameters to help improve predictions for our scenario

Implementation: Trading Signal

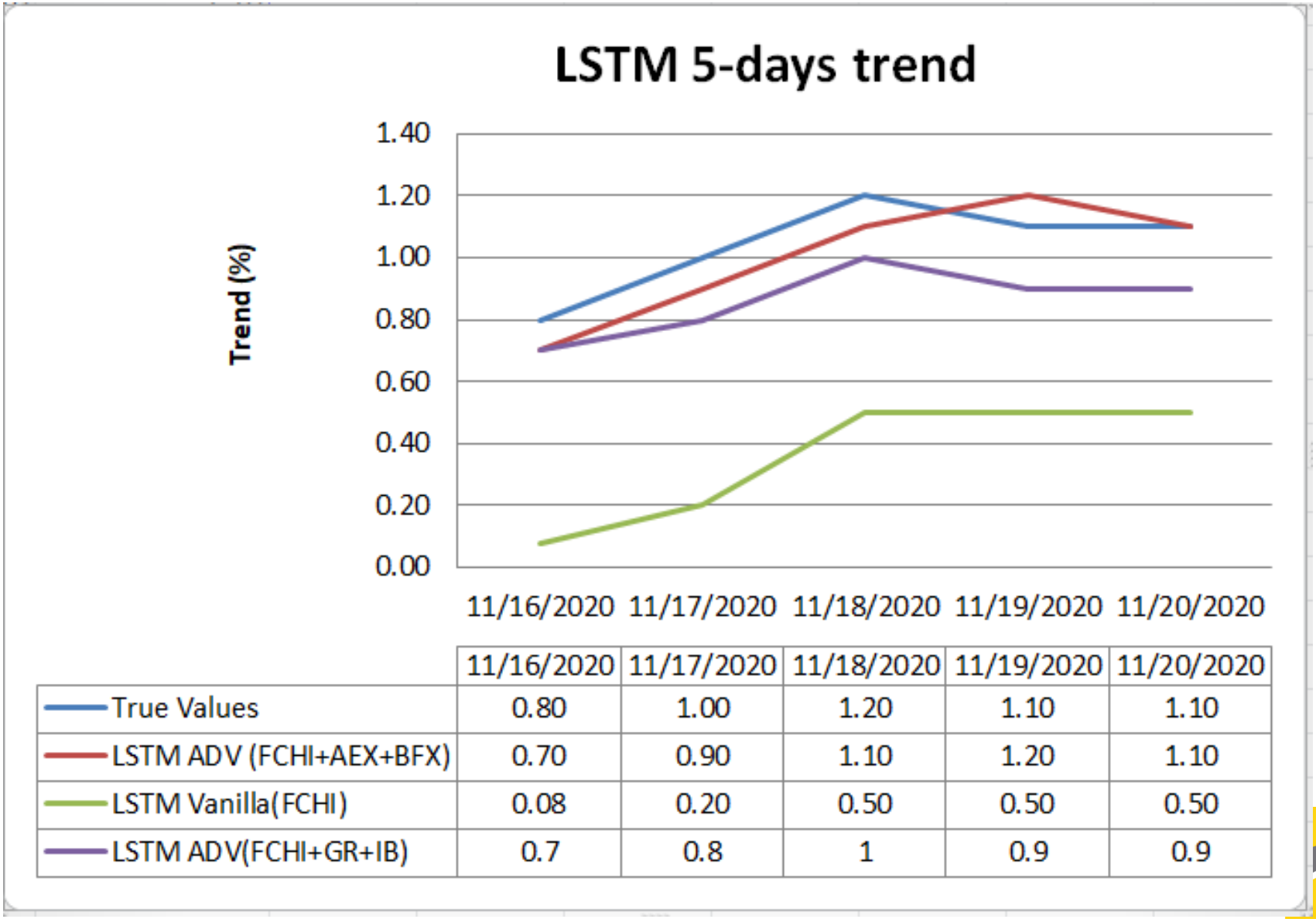
- Five trading signals that perform a weighted average conclusion
- Conclusion reflects the trend of the data (uptrend, downtrend or sideways)
- Backtests show that long-term predictions are more certain

Part 2 : LSTM

Context

- LSTM is a deep neural network model used for stock prediction
- Contrary to SVM, it is a regression model for the near future predictions
- Training datasets that are used for LSTM
  1. **FCHI** - CAC40 French stock market index
  2. **AEX** - Amsterdam Stock Exchange
  3. **BFX** - BEL20 Belgium Stock Exchange
  4. **GR** - German Stock Market
  5. **IB** - Spanish Stock Market

LSTM results



	11/16/2020	11/17/2020	11/18/2020	11/19/2020	11/20/2020
True Values	0.80	1.00	1.20	1.10	1.10
LSTM ADV (FCHI+AEX+BFX)	0.70	0.90	1.10	1.20	1.10
LSTM Vanilla(FCHI)	0.08	0.20	0.50	0.50	0.50
LSTM ADV(FCHI+GR+IB)	0.7	0.8	1	0.9	0.9

Credit risk Evaluation

Context and Methodology

Objective

Automation of the process of approving an overdraft.

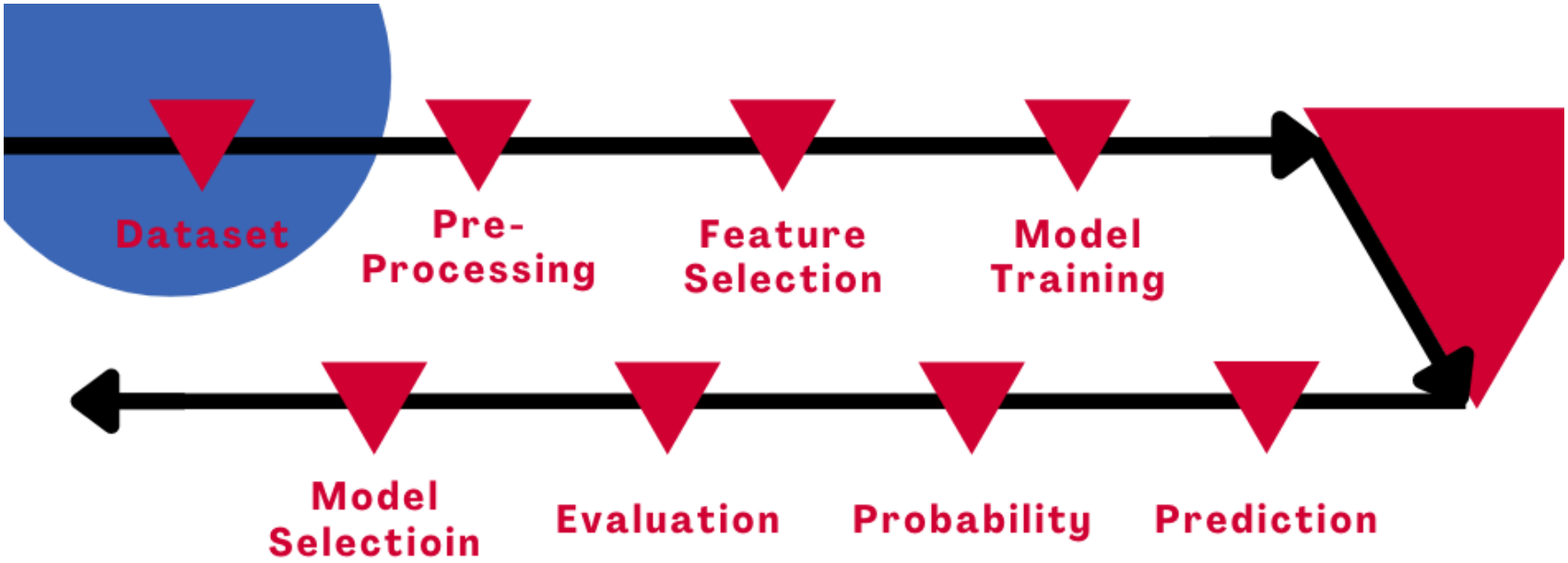


Figure: Schema of methodology

Final model & evaluation

- Balanced Random Forest
- Simple model and similar results to other methods

Method	Specificity	Sensitivity	Accuracy	AUC
BRF	0.785	0.707	0.712	0.746

Model explainability

- The banks has to justify the decision of approving or not approving the credit
- We calculate the most important features for each prediction separately

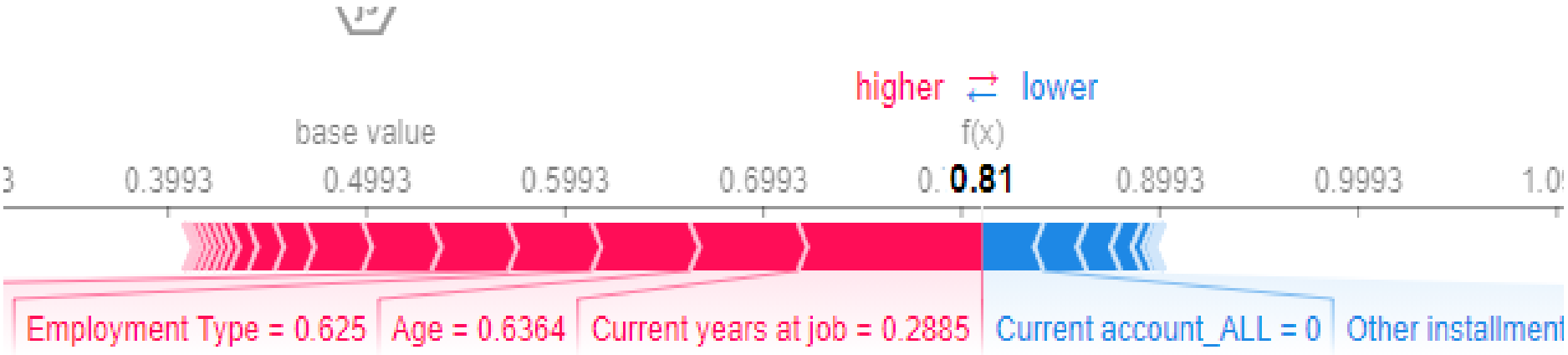


Figure: Local feature importance

Dataset exploration

The goal of the second part was to develop a general credit assessment tool

We explored a few datasets :

- UCI Repository : Mutiple credit datasets from various countries, only a small number of usable features
- PAKDD 2010 : Too few usable features
- Home Credit Default Risk : Unusable in our context
- Fast Credit : Too few samples

Lending Club dataset

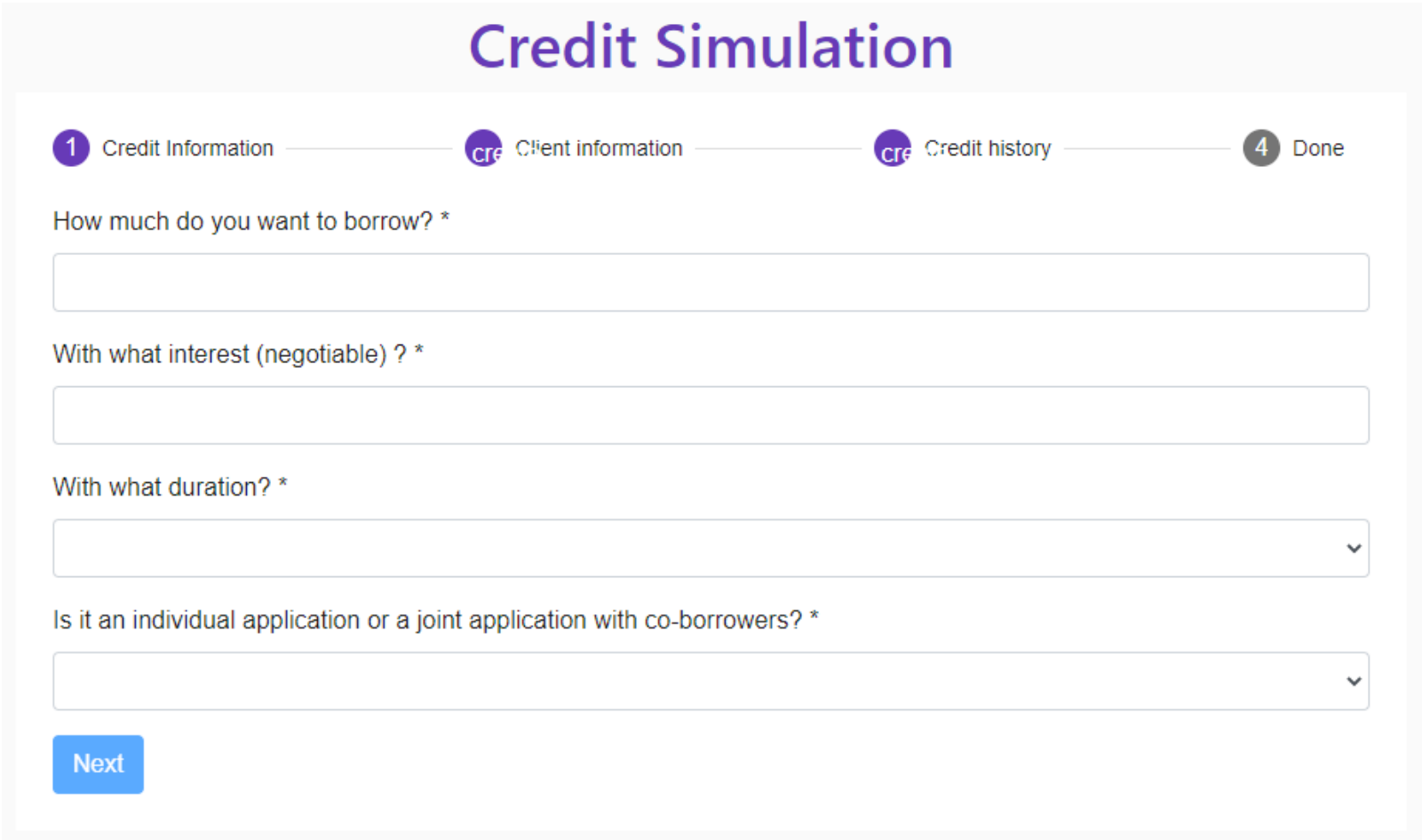
Huge number of samples (> 1M), sufficient number of usable features, the only dataset we found really usable for our purpose.

Model generalization

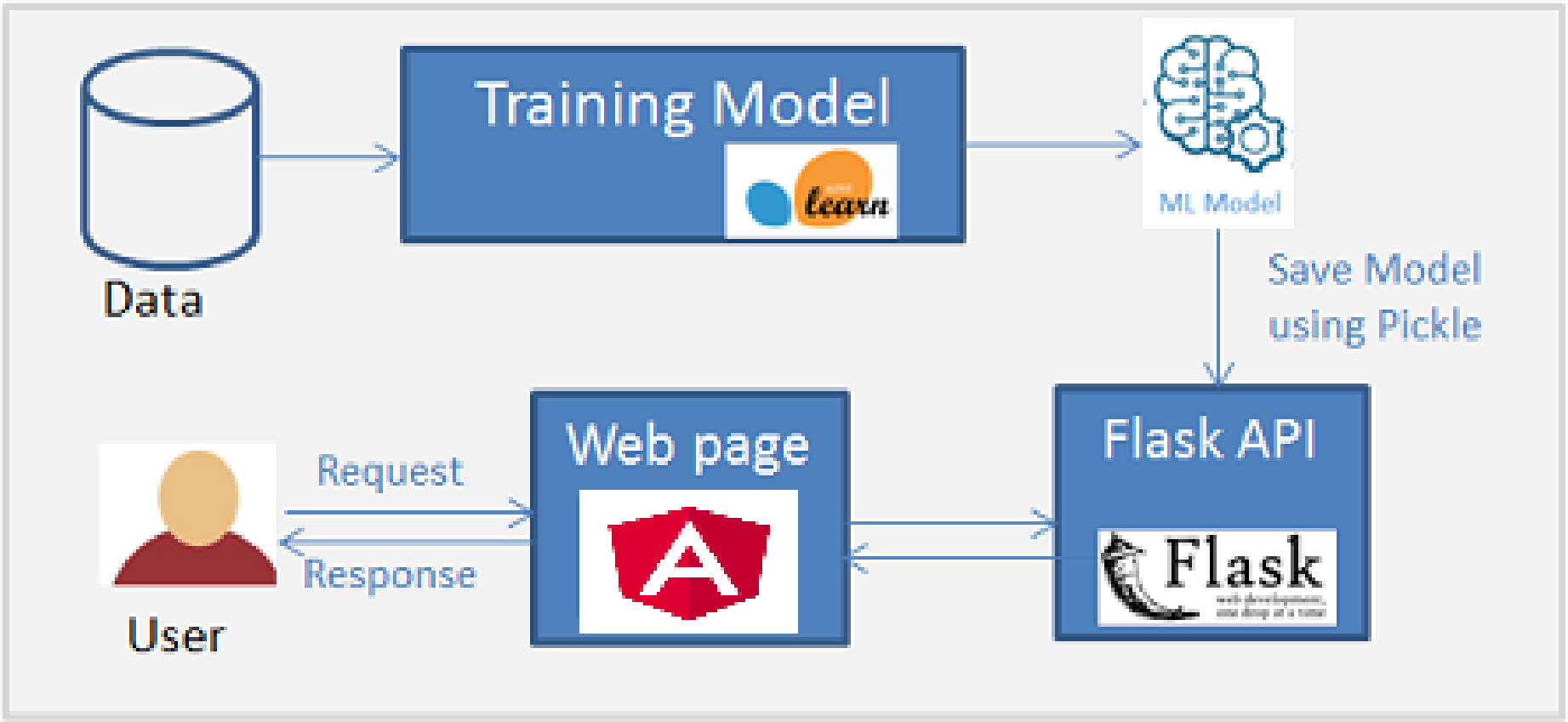
- We finally selected 13 features from the generalization Dataset for the final model analysis.
- From 3 different aspects :
  - Credit information required by the client
  - Client's Personal information
  - Traceability of the past credit demanded by the client
- Balanced Random Forest Model with AUC = 0.64

Web application functionalities

- Require all the necessary information from the client using a multi step form
- Output the probability that the client gets an approbation for the credit



Web application architecture



Project before & after

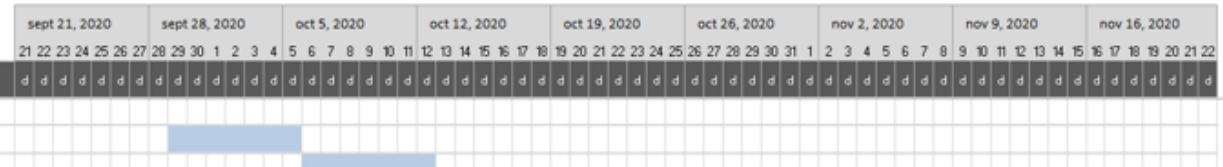
Comparing Gantt charts I

What use of AI in finance ?

Project Start: 22/09/2020

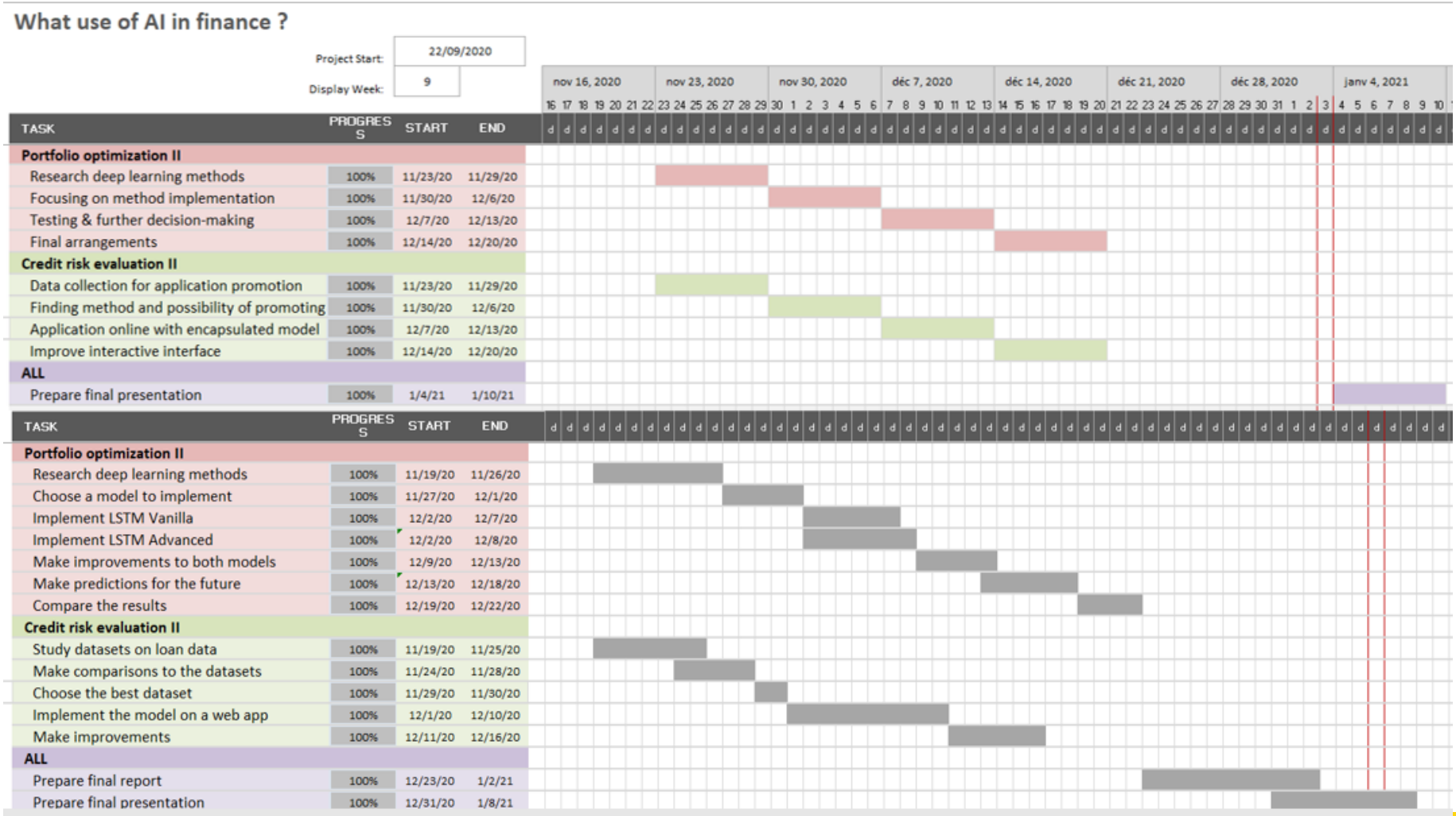
Display Week: 1

TASK	PROGRESS	START	END
ML Applications Estia			
Gather information about ML apps	100%	9/29/20	10/5/20
Prepare presentation	100%	10/6/20	10/12/20
Portfolio optimization I			
Researching financial indicators	100%	9/29/20	10/5/20
Study ML algorithms & application interface	100%	10/6/20	10/12/20
Check current implementation & dataset req	100%	10/13/20	10/19/20
Research on datasets compatible with requir	100%	10/20/20	10/26/20
Integrate in the French stock market	100%	11/2/20	11/8/20
Credit risk evaluation I			
Data analysis	100%	9/29/20	10/5/20
Model trial	100%	10/6/20	10/12/20
Researching and reading paper	100%	10/13/20	10/19/20
Improvement with different methods	100%	10/20/20	10/26/20
Model packaging	100%	11/2/20	11/8/20
ALL			
Paper study	100%	9/22/20	9/28/20
Preparation for presentation	100%	11/9/20	11/15/20
Midterm presentation	100%	11/16/20	11/22/20



TASK	ASSIGNED TO	PROGRESS	START	END
ML Applications	Estia			
Gather information about ML apps		100%	9/28/20	10/5/20
Prepare presentation		100%	10/5/20	10/12/20
Prepare report		100%	10/13/20	10/20/20
Portfolio optimization I				
Researching financial indicators		100%	9/28/20	10/5/20
Study ML algorithms (SVM particularly)		100%	10/6/20	10/11/20
Check current implementation		100%	10/12/20	10/14/20
Test & implement SVM model		100%	10/13/20	10/21/20
Integrate in the French stock market		100%	10/20/20	10/25/20
Credit risk evaluation I				
Analyze overdraft dataset		100%	9/28/20	10/6/20
Choose the most accurate model		100%	10/6/20	10/11/20
Improve the model		100%	10/12/20	10/17/20
Model packaging		100%	10/16/20	10/21/20
Explain results using SHAP method		100%	10/21/20	10/25/20
ALL				
Paper study		100%	9/22/20	9/28/20
Prepare presentation		100%	11/11/20	11/18/20

Comparing Gantt charts II



State of the project

Portfolio Optimization

- Start mostly from scratch
- Based on Haifei Zhang's Master Thesis
- Adapting to CAC40 dataset
- Improving SVM model
- Adding Deep Learning methods (LSTM)
- Make 7-day predictions

Credit Risk

- Start mostly from scratch
- Based on Zinan Zhou's Master Thesis
- Implement Balanced Random Forest for Overdraft
- Adapt to Lending club dataset
- Build a web application

Conclusions

- The study of AI applications helped to create an idea of how the AI and finance world are connected as well as how our two applications are placed.
- For the portfolio management problem, we studied previous and state-of-the-art algorithms and implemented for the French stock market dataset. We used different models of SVM and LSTM.
- For the credit risk assessment problem, we analysed the loan data and after implementing various methods we concluded that the best method to use is **Balanced Random Forest**. We chose to implement it to the Lending club dataset and built a web application.