**NETCOMPANY – PROJECT DATA ANALYSIS (PDA)**

**PROJECT PROPOSAL**

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

.

|  |  |
| --- | --- |
| **Version:** | 1.4 |
| **Status:** | Not Approved |
| **Approver:** | Ngô Thái Bình  Nguyễn Thị Diễm Trang |
| **Author:** | Nguyễn Bảo Nguyên  Quách Hoàng Minh  Ngô Gia Hân  Nguyễn Vũ Anh Thư |

|  |  |  |  |
| --- | --- | --- | --- |
| **Version** | **Date** | **Authors** | **Rational** |
| 1.0 | October 22th, 2021 | Nguyễn Bảo Nguyên  Quách Hoàng Minh  Ngô Gia Hân  Nguyễn Vũ Anh Thư | Outline Proposal |
| 1.1 | October 23th, 2021 | Nguyễn Bảo Nguyên | First Draft |
| 1.2 | October 29th, 2021 | Nguyễn Bảo Nguyên  Quách Hoàng Minh  Ngô Gia Hân  Nguyễn Vũ Anh Thư | Update Scope Objective, Summary Milestone Schedule, Scope, Summary Budget |
| 1.3 | November 2th, 2021 | Nguyễn Bảo Nguyên  Quách Hoàng Minh | Update Scope Objective, Milestone Schedule, |
| 1.4 | November 11th, 2021 | Nguyễn Bảo Nguyên  Quách Hoàng Minh  Ngô Gia Hân  Nguyễn Vũ Anh Thư | Update all sections |

Table of Contents

[1. Glossary 4](#_Toc87877900)

[2. Foreword 5](#_Toc87877901)

[3. Overview 6](#_Toc87877902)

[3.1 The document 6](#_Toc87877903)

[3.2 The client 6](#_Toc87877904)

[3.3 The Problem and Solution 6](#_Toc87877905)

[4. Scope and objectives 8](#_Toc87877906)

[4.1 Project Objectives and Success Criteria 8](#_Toc87877907)

[5. Scope 9](#_Toc87877908)

[ Constraints 9](#_Toc87877909)

[5.1 Requirements 10](#_Toc87877910)

[5.2 Project Deliverables 11](#_Toc87877911)

[6. Project methodology and Approach 12](#_Toc87877912)

[6.1 Vision 12](#_Toc87877913)

[6.2 Method and Approach 12](#_Toc87877914)

[ Scrum 12](#_Toc87877915)

[ Project Management 12](#_Toc87877916)

[7. Risks Management 13](#_Toc87877917)

[8. Skill Analysis 15](#_Toc87877918)

[8.1 Skill needs for the project 15](#_Toc87877919)

[8.2 Skill matrix 15](#_Toc87877920)

[9. Summary Milestone Schedule 16](#_Toc87877921)

[9.1 Gantt Chart 16](#_Toc87877922)

[10. Summary Budget 19](#_Toc87877923)

[11. Project Approval Requirements 20](#_Toc87877924)

[12. Project Manager 20](#_Toc87877925)

[13. Disclaimer 21](#_Toc87877926)

[14. References 22](#_Toc87877927)

# Glossary

|  |  |
| --- | --- |
| **Glossary** | **Explanation** |
| API | Application Programming Interface |
| MVP | Minimum Viable Product |
| PDA | Project Data Analysis |
| PI | Program Increment |
| R&D | Research and development course |
| UAT | User Acceptance Testing |

# Foreword

This is the proposal report for an interesting Research and Development project. This project will give an effective web site for the investors to do a lot of market activities in just one place. It will give not only the stock price but also the technical anticipated by the most state-of-the-art Artificial Intelligence and Machine Learning algorithms. Not only that, but the website will also provide the reliable stocks’ new, specific instructions for the beginners and moreover.  
We will provide a friendly-for-user website about the Vietnam stock market for the investors with the easiest instructions for the beginner to join in Vietnam’s stock market.

This document will show the potential of the project as well as the ability of the development team to meet the requirements of the customer. The document will be provided for the clients Mr. Binh and Ms. Trang and the professors are responsible for the R&D course.   
We would like to express our sincere thanks for the team members who joined in this project, Mr. Ngo Thai Binh and Ms. Nguyen Thi Diem Trang from Netcompany, the two professors Dr Lam Quang Vu, MSc. Petteri Kaskenpalo and our supervisor Dr. Le Thi Nhan.

# Overview

## The document

This is the project proposal document for the Data analysis project provided by Netcompany for a group of four students in the Research and Development course taught by University of Science and Auckland university of Technology.

This Data analysis project gave the initial description as project management data analysis at first. However, in the process of exchanging and interacting with the customer, the project requirements turn into stock market predictions based on available data sets. The customer wants this project to be a fixed price and fixed duration project, and the final product will be available at the end of April 2022.

The risks that could cause the lack of working progress will be identified and provided the solution of each risk with the trigger might cause them.

The skill analysis of each team member will be provided with the specific level of each skill.

The estimated budget for this project will be 3000$ and this price will not change though all the working progress.

And this proposal is written to implement projects and solve problems for customer.

## The client

Netcompany is a Danish IT consultancy headquartered in Copenhagen, Denmark. The company was founded in 2000 by André Rogaczewski, Claus Jørgensen and Carsten Gomard and since that, it has become an international company with more than 3000 staff in its offices in Denmark, United Kingdom, Norway, The Netherlands, Poland, and Vietnam.

## The Problem and Solution

Netcompany has a group of investors that wants to have a website to check the stock market’s information. For convenience in grasping market trends, customers want an application to view the stock market price and provide reliable news about the stock market. The website should also give the auto technology prediction in the coming hours or days without hiring stock forecasters.   
However, this will be a difficult problem for software developers because the stock market is complex and has many different price ranges. In addition, the stock price is always changing, and it fluctuates continuously from time to time.

The stock market is a volatile place. The cause of that fluctuation comes from many sources: market trends, investor sentiment, company's financial statements, etc. There are two ways to analyze securities: Fundamental analysis and technical analysis.

Fundamental analysis assumes that stock prices do not inevitably reflect the true intrinsic value of the underlying business. And our solution for our customers lies in the second analytical method. Technical analysis generally assumes that a stock's price reflects all available information and that prices generally move according to trends. In other words, by analyzing a stock's price history, you may be able to predict its future price behavior. If you have ever seen someone trying to identify patterns in stock charts or discussing moving averages, that's a form of technical analysis (Frankel, 2021).

The major goal of this project is to create a stock predicted website. We shall make an online website have an online stock market price list with a service for the customer registered account. Beside the main page, we also have the technical analysis chart that will predict the future trend of stock by AI algorithm. At the end of our project, we will have a web page to show the table of stock information, study pages for beginners, account protection, information about stocks pages and the charts of real price and predicted price, with some other interesting functions.

# Scope and objectives

## Project Objectives and Success Criteria

The purpose of this PDA project is to create a website that displays full stock market information, from which customers can refer to make decisions to buy or sell stocks. The available data sets will be analyzed and predicted through the application of the most suitable algorithm. Based on those predictions, along with additional information that the customer knows about, the customer can decide whether to invest in that stock or not. The business objective of this project directly assists clients in predicting stock prices in the coming hours or days.

Mutual goals for the milestones and deliverables for this project have been identified. To be successful in this PDA project, the following objectives must be met within the specified time and budget allocation:

* Complete a web app that meets the requirements within the next 6 months
* Finalize product related documents within the next 6 months
* Complete the chosen algorithm within the next 4 months.
* Create demo products within the next 4 months.
* Archive versions of the application for the next 6 months.
* Prepare all presentation materials for the next 6 months to present to stakeholders.
* The dataset must come from a reliable source and the dataset will be obtained from the Vietnam’s stock market API. We will fetch data from the API for our research and testing. Therefore, the dataset can be always updated and add-in. We found some potential API such as Vietstock API, Vndirect API, which have a lot of stocks from some big companies.

We will apply the appropriate algorithmic model found to predict the future by hours and days instead of predicting for months or years. The purpose of this is to reduce most of the risk of the effects of unexpected factors outside the world economic market (Wang, 2012).

Accept criteria of the PDA project include:

* web app that shows charts representing stock market prices as well as with the table of the stock information.
* The algorithm can predict about 60% correctly.
* Successfully deploying web apps and functions to web browsers.
* Price chart viewing and price prediction functions can work well.
* Have an account to meet the account protection requirement of users.
* It is possible to switch back and forth between different types of shares

# Scope

### Constraints

* Costs must not exceed the estimated budget.
* The final product must be released within 6 months.
* Must have a UAT version for customers to try before releasing the final version.
* The final product must satisfy the Minimum Viable Product (MVP).
* There are documents and resources that prove the predictions are reliable.

## Requirements

|  |  |  |
| --- | --- | --- |
| **ID** | **Feature** | **Description** |
| 1 | Register account | As a customer, I'm able to register account |
| 2 | Login/Logout | As a customer, I'm able to login/logout |
| 3 | Overview page | As a customer, after logged in successfully, I'm able to see the overview page:                  + I can see the search button for search index                  + I can see the list of capitalization-weighted index of all companies listed on the stock exchange and their values (date, point, volume)). E.g.: (these indices are just an example. We will choose them depend on what dataset we can find)                                  + VNIndex of HoSE                                  + VN30                                  + HNX-Index                                  + HNX30-Index                                  + UPCOM                  + Favorite indices list:                                  + If it is unavailable, display empty                                  + I can see the information of each stock index                                                  + date, price, volume                                                  +I can click the index to view the chart |
| 4 | Search index | As a customer, I'm able to search index  + I can search by index  + I can add index search result to favorite list  + I can click the index search result to view its chart |
| 5 | Index chart | As a customer, I'm able to view index chart:                  + I can see the list of charts that I would like to view. They are line chart (by default) & candlestick chart                  + I can view the index chart:                                  + line chart:                                                  + x-axis: time unit (date or month)                                                  + y-axis: price (VND)                                                  + prediction area                                  + When I hover to the line, I can see the popup which will display a price                  + I can choose to candlestick chart to view index                  + I can see the bar chart which will display the volume of index by time unit in the bottom                                  + I can see the prediction volume area                                  + When I hover to the column, I can see the popup which will display a price                  + I can do comparison by choosing other indices                                  + In the current chart, I can see one more line chart of the comparing index. |
| 6 | Setting reminder | As a customer, I can set reminder:  + I can choose the expectation point in the prediction area in a chart  + I can create reminder by adding title, timing, content.  + I can receive the reminder via email. |
| 7 | Manage user profile | As a customer, I'm able to view and update my profile:                  + I can view and update username, email, avatar, password                  + I can view reminder lists                                  + I'm able to remove a reminder in a list |

Additional requirements may be added as needed, with the approval of the project sponsor, as the project progresses.

## Project Deliverables

According to each PI, there will be documents that need to be completed accompanied by some parts of the overall product.

* First PI:
* A0100 - Analysis Report
* O0500 - Software Architecture - Clarification Phase
* Demo and Compare Algorithms
* Second PI:
* D0100 - User Interface Guidelines
* DD130 - Detailed Design
* D0160 - User Interface Design
* O0500 - Software Architecture
* Mockup
* Algorithm Report
* Third PI:
* Manual Test Document
* Demo Product
* Fourth PI:
* Final Product

# Project methodology and Approach

## Vision

To deliver deliverable products, we will research and identify the most suitable Machine Learning algorithm to predict the stock market based on collected data sets. After that, we will build a web application to represent the prediction results. The project is divided into four project increments.

## Method and Approach

### Scrum

Scrum has been used since the early 1990s as a framework for developing, providing, and maintaining complex products. Scrum illustrates the relative effectiveness of the product management and work techniques, allowing to make improvements to the product, team, and workplace environment.

Therefore, we will apply the Scrum process to manage this PDA project. We will divide the project into four main PIs and each PI will be broken down into smaller appropriate Sprints. Before each Sprint, there will be a planning meeting to divide tasks and have weekly status reports. On average, each sprint will last about 2 weeks.

### Project Management

Jira will be used as a project management system for the PDA project, including task division, task assignment, and performance tracking, time tracking. Beside that, we also use ClickUp to draw Gantt Chart.

# Risks Management

The following risks for the PDA project have been identified. The project manager will determine and employ the necessary risk mitigation/avoidance strategies as appropriate to minimize the likelihood of these risks: (1-lowest, 4-highest)

* **Probability**: from 1 (lowest) to 4 (highest)
* **Impact**: from 1 (lowest) to 4 (highest)
* **Risk Score**: ***Probability x Impact***
* **Risk ranking**: based on risk score
  + I (1, 2): Not serious, do not need to be immediately resolved.
  + II (3, 4): Not serious, need to be immediately resolved.
  + III (6, 8, 9): Serious, do not need to be immediately resolved
  + IV (9, 12, 16): Serious, need to be immediately resolved

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Risk Identification | | Qualitative Rating | | | | Risk Response | |
| Risk | **Risk Category** | **Probability** | **Impact** | **Risk Score** | **Risk Ranking** | **Risk Response** | **Trigger** |
| Human absence due to personal issues | Human Resources | 2 | 3 | 6 | III | Reschedule the tasks or assign them to another member | Team members must inform all team. |
| The data set is not large enough to affect the accuracy of the algorithm | Technical Accuracy | 1 | 4 | 4 | II | 1. Change another appropriate dataset.  2. Add more data to the dataset. | The accuracy of result prediction fluctuates throughout many runs. |
| Poor quality of data | Technical Accuracy | 2 | 4 | 8 | III | 1. Change to another appropriate data set.  2. Cleaning the dataset. | There are some duplicated data, or missing value. |
| The accuracy of the prediction algorithm is less than 60% | Technical Accuracy | 2 | 4 | 8 | III | Change another algorithm to test if the accuracy is higher | Throughout 5 runs, the accuracy results all less than 60% |

In addition, risks related to people (conflicts, ...), technical (lack of skills, knowledge, …), in terms of time (time consuming tasks, ...) will be resolved in Scrum retrospective at each sprint among team members.

# Skill Analysis

## Skill needs for the projects

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Group | Skill Expertise | Level | | | | |
| **1** | **2** | **3** | **4** | **5** |
| Technical Skill | Coding skill |  |  | x |  |  |
| React |  | x |  |  |  |
| MongoDB |  | x |  |  |  |
| NumPy, Pandas, Matplotlib |  | x |  |  |  |
| Scikit-learn, TensorFlow | x |  |  |  |  |
| Team Skill | Presentation |  |  |  | x |  |
| Research |  |  |  | x |  |
| Problem Solving |  |  |  | x |  |
| Conflict management |  |  | x |  |  |
| Time managing |  |  |  |  | x |

Our team is good at presentations, research and problem solving. We are optimistic about our time management skills, too. Although we have fundamentals of coding, our skills in using Python libraries are not very good. Furthermore, we need to have domain knowledge in the stock market. Therefore, our team will improve knowledge, skills and do the tasks in the project simultaneously.

## Skill matrix

A screenshot of a computer

Description automatically generated with medium confidence

# Summary Milestone Schedule

The project Summary Milestone Schedule is presented below.  As requirements are more clearly defined this schedule may be modified.  Any changes will be communicated through project status meetings by the project manager.

|  |  |  |
| --- | --- | --- |
| Project Milestone | Start Date | Due Date |
| Project Start | October 14th, 2021 | October 21st, 2021 |
| PI Planning | October 22nd, 2021 | November 5th, 2021 |
| First PI | November 6th, 2021 | December 31st, 2021 |
| Second PI | January 3rd, 2022 | February 28th, 2022 |
| Third PI | March 1st, 2022 | April 1st, 2022 |
| Final PI | April 4th, 2022 | April 30th, 2022 |
| Project Complete | - | April 30th, 2022 |

## Gantt Chart

Below is a Gantt chart showing the major tasks of the team. However, for each sprint, these tasks will be broken down accordingly and be assigned to each member.

Timeline

Description automatically generated

Figure 1: PI Planning from 22/10/2021 to 5/11/2021

Timeline

Description automatically generated

Figure 2: PI 1 from 8/11/2021 to 31/12/2021

Graphical user interface

Description automatically generated with medium confidenceFigure Figure 3: PI 2 from 3/1/2022 to 28/2/2022

Week from 31/1/2022 to 4/2/2022 is Lunar New Year holidays

Timeline

Description automatically generated with medium confidence

Figure 4: PI 3 from 1/3/2022 to 1/4/2022

Timeline

Description automatically generated

Figure 5: PI 4 from 4/4/2022 to 29/4/2022

# Summary Budget

Our project is a fixed cost and fixed duration project.

The following table contains a summary budget based on the planned cost components and estimated costs required for the successful completion of the project.

|  |  |
| --- | --- |
| Project Component | Component Cost |
| Personnel salary | $2,400 |
| Traveling fee | $96 |
| Other (license...) | $500 |
| Total | $3,000 |

* Number of people: 04
* Price 1 person/hour: $2
* Time/day: 2,5 hours/day
* Personnel salary: 1200 (hours) x $2 (person/hour) = $2,400
  + Total Estimated time: 4 (person) x 2,5 (hours/day) x 120 (days) = 1200 (hours).
* Traveling fee: $2 \* 2 (every 2 weeks) \* 6 (months) \* 4 (person) = $96

# Project Approval Requirements

The complete project is defined when a product that predicts stock market trend is tested and proven to be the best solution by the accompanying research papers. Along with that, all technical and process documents are handed over to the company, the product must be fully deployed within the time and cost constraints indicated in the proposal. In addition, this successful measure needs to include instructions for using and installing the product along with a video showing the application process in action. Success will be supported by the project sponsor Netcompany, specifically the representative of Miss. Nguyen Thi Diem Trang was identified.

# Project Manager

Ngo Thai Binh is the project manager, and he is responsible for creating main tasks and deadlines for each task. This task is given to the team leader, Nguyen Bao Nguyen. Nguyen divides the tasks into 4 PI and identifies the smaller tasks of each team member in the PI. Mr. Binh’s team has four members including a team leader, business analyst, developer, and tester. After each phase of the project, the product will be delivered by Mr. Binh to the Project Owner, Nguyen Thi Diem Trang.

# Disclaimer

**Clients should note the general basis upon which the Auckland University of Technology undertakes its student projects on behalf of external sponsors:**

*While all due care and diligence will be expected to be taken by the students, (acting in software development, research or other IT professional capacities), and the Auckland University of Technology, and student efforts will be supervised by experienced AUT lecturers, it must be recognized that these projects are undertaken in the course of student instruction. There is therefore no guarantee that students will succeed in their efforts.*

*This inherently means that the client assumes a degree of risk. This is part of an arrangement, which is intended to be of mutual benefit. On completion of the project, it is hoped that the client will receive a professionally documented and soundly constructed working software application, some part thereof, or other appropriate set of IT artefacts, while the students are exposed to live external environments and problems, in a realistic project and customer context.*

*In consequence of the above, the students, acting in their assigned professional capacities and the Auckland University of Technology, disclaim responsibility and offer no warranty in respect of the “technology solution” or services delivered, (e.g. a “software application” and its associated documentation), both in relation to their use and results from their use.*

# References

*Stock Market Data (NASDAQ, NYSE, S&P500)*. (n.d.). (Kaggle) Retrieved from https://www.kaggle.com/paultimothymooney/stock-market-data