CryptOL™ Project Design Document

CS-492A

Group: CryptOL™

Alexander Gribtsov, Chris Woszczak, John Guseman

**Introduction:**

Implementation of different types of artificial intelligence (AI) algorithms can produce fit data that predicts the price of different cryptocurrency types such as BitCoin. Project CryptOL™ uses a web-based application that allows us to integrate multiple AI models for cross comparison, the quality of each model and the accuracy of their predictions is coveted. Project CryptOL™ will forecast the price of several cryptocurrencies 15 minutes into the future. CryptOL™ is a full stack web based software that features a graphical user interface (GUI).

**System Overview:**

Users will select a cryptocurrency to forecast from the GUI. The GUI will chart the current price of the cryptocurrency alongside its forecasted price. The framework of the application will be designed in Python using Scikit and TensorFlow libraries on the Anvil architecture. In the future, multiple algorithms will use parts of each currently functioning prediction algorithm in synchrony.

**Design Considerations:**

Display of real time data is extremely important to a prospective user of our software. Issues that were addressed before the design phase began:

* Assumptions and Dependencies: the project depends heavily on AI libraries such as Scikit and TensorFlow
* General Constraints: the libraries are executed on the server side of the application for security purposes sometimes causing delays of data relay. Internet connection speed is another considerable constraint
* Goals and Guidelines: The goal of the software is to output data prediction using multiple algorithms that is always with in 25% error fit
* Development Methods: Anvil architecture provides an IDE in which the majority of the code is designed. The primary code language is Python 3.0

**Architectural Strategies:**

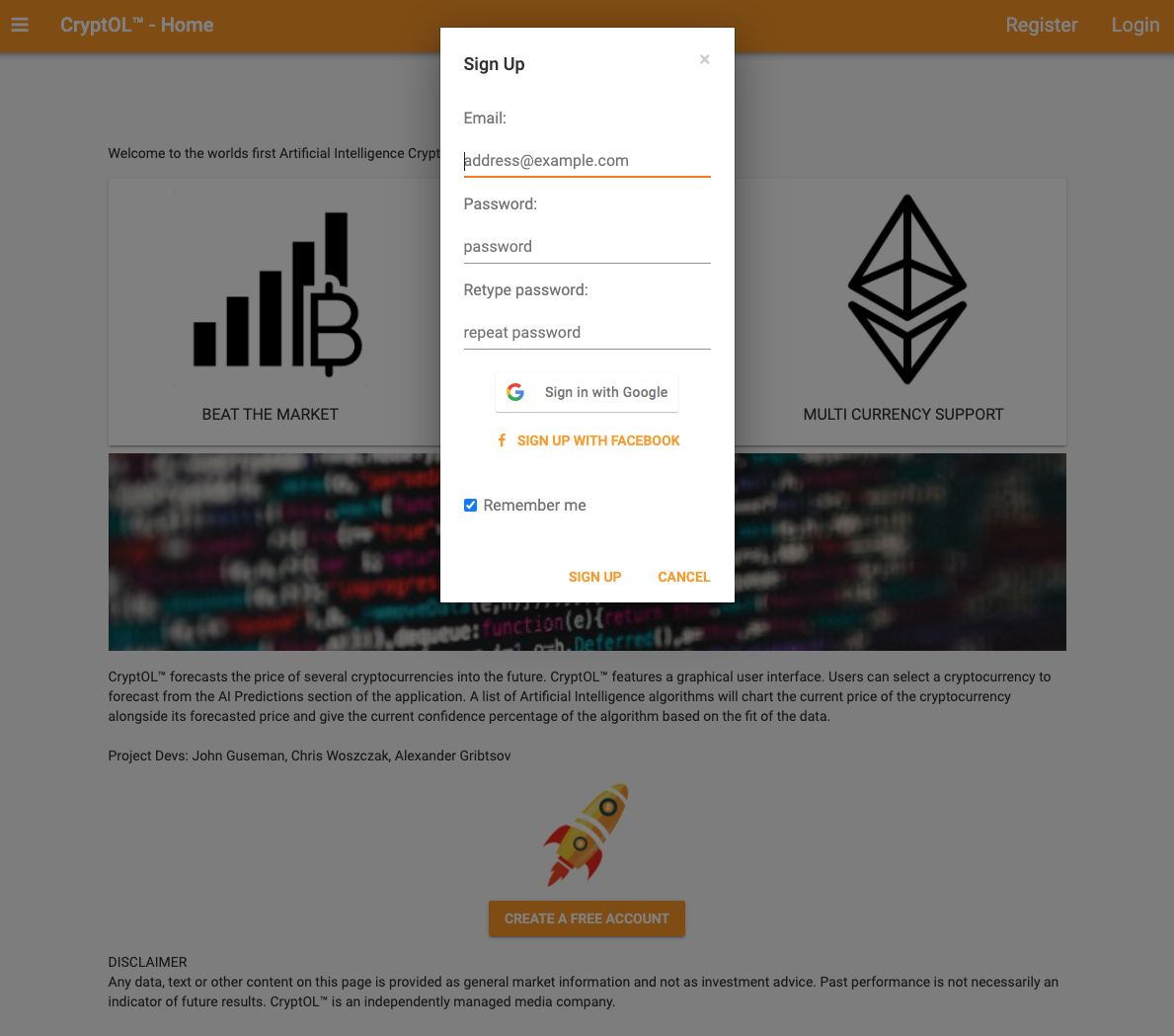
The main strategy employed in the software is to pull live data from a currency market API and manipulate the data pushing it through the algorithms designed based on machine learning using Keras and TensorFlow. We employed several techniques and algorithms for displaying data to the user which included: linear regression, logistic regression, long short-term memory (LSTM), time series, survival modeling, online machine learning, neural networks.

**System Architecture:**

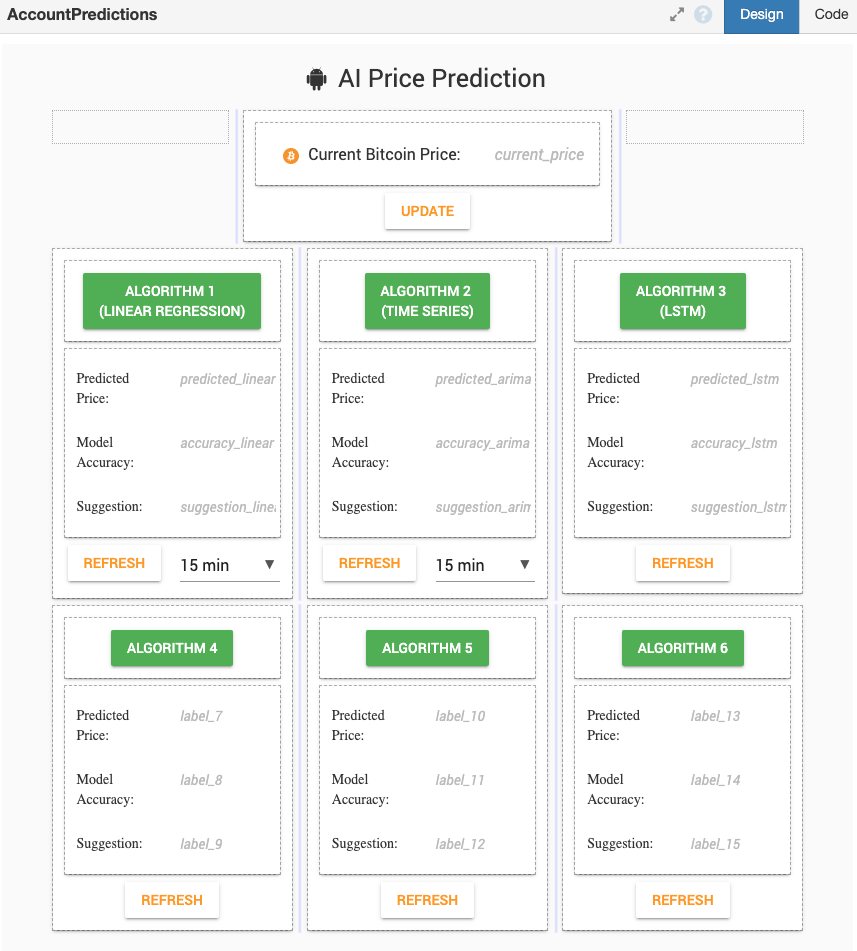
Everything is coded in Python and integrated into the Anvil IDE. The IDE integrates the client and server side functions with an internal database for user credentials as well as other methods employed throughout the program. On the server side module we import and activate the Keras and TensorFlow library packages, we then have several modules on the server side that utilize the packages to train data imported from yahoo finance and split into two sets, the training set and the test set. Although this is the structure of our Logistic regression algorithm output, other algorithms such as LSTM follow a different design structure. In general we designed four working algorithms each with a unique training and testing technique.

**GUI:**

The graphical user interface is designed in a classic home landing setting with an immediate request for registration as seen in the display below:

****

While there are many instance interfaces in the software the main focus of the application is to display the currently (6) currency prediction algorithm results at a given notice as seen in the internal diagram generated by the Anvil IDE below:

****

The page that displayed this desirable data can only be accessed by registered users. The first algorithm data is available to all registered users. To gain access to the remaining 5 algorithms the user must upgrade to a paid premium version of the application.

**Detailed System Design:**

Project CryptOL is a web based full stack application. The design, implementation and deployment of this project is covered in detail. Six algorithms were developed for the user display.

**Algorithms and their details**

**Linear Regression Algorithm**

In statistics, linear regression is a linear approach to modelling the relationship between a scalar response and one or more explanatory variables (also known as dependent and independent variables).

In CryptOL™ linear regression is used to extrapolate a trend from Yahoo Finance library assets. Linear regression and ordinary least squares (OLS) are decades-old statistical techniques that can be incorporated into machine learning to extrapolate a trend in the observed asset and predict the direction of future price movement.

**Time Series**

Time series forecasting is used to predict future values based on previously observed values and one of the best tools for trend analysis and future prediction.

A famous and widely used forecasting method for time-series prediction is the Auto-Regressive Integrated Moving Average (ARIMA) model. CryptOL™ ARIMA model is capable of capturing a suite of different standard temporal structures in time-series data.

**Long Short-Term Memory (LSTM)**

LSTM is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections.

In CryptOL™ LSTM is very powerful in sequence prediction problems because it is able to store past information. This is important in our case because the previous price of BitCoin is crucial in predicting its future price.

**Logistic Regression**

Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist.

In CryptOL™ Logistic Regression algorithm is a part of the Supervised Learning method of Machine Learning. It is a statistical method for the analysis of a dataset. It has one or more independent variables that determine an outcome. We use a training set and a test set of data to predict the selling price of BitCoin.

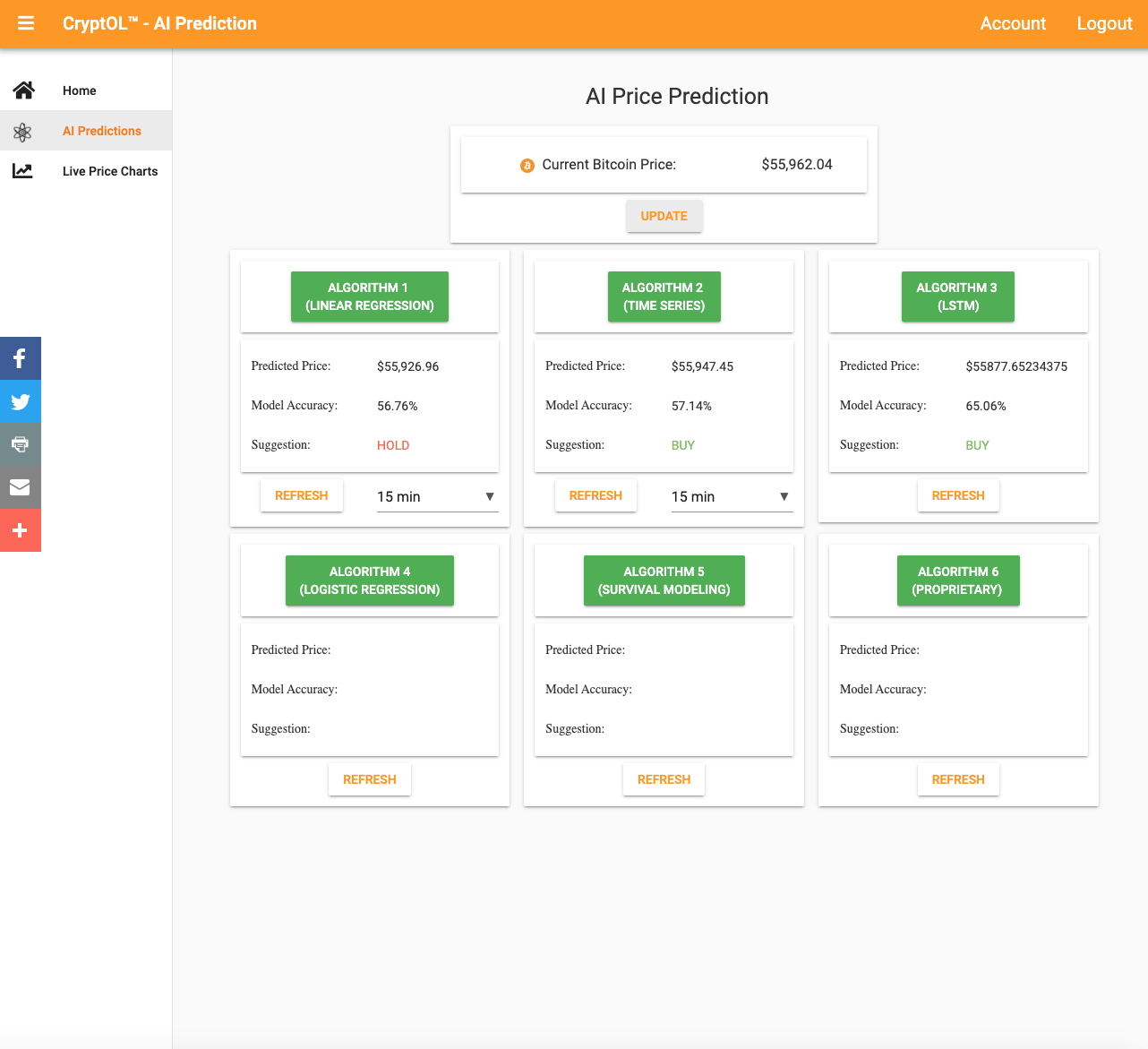
**Survival Modeling**

Survival analysis is a statistical method that aims to predict the time to an event, such as death, the diagnosis of a disease or the failure of a mechanical part.

CryptOL™ adopts Cox's hazard model to predict BitCoin’s future rising or dropping probabilities. Specifically, we define the problem of Buy-and-Sell-Point Prediction from the survival analysis perspective. We apply the trained model for the cryptocurrency market forecasting on two cryptocurrencies traded on the CoinBase exchange.

**CryptOL™ Proprietary Algorithm**

Our “secret” proprietary algorithm incorporates the best in prediction techniques of all the standardized algorithms combined to produce the most accurate comprehensive value that is possible through Machine Learning.



**Glossary:**

Artificial intelligence (AI) - refers to the simulation of human intelligence in machines that are programmed to think like humans and mimic their actions.

Anvil - Proprietary web application full stack development framework

BitCoin - The original most popular cryptocurrency

Full Stack Application - Software that combines and employes, client side code, server side code, html CSS and database(s)

Graphical user interface (GUI) - a visual way of interacting with a computer using items such as windows, icons, and menus, used by most modern operating systems.

Integrated development environment - (IDE) - is a software application that provides comprehensive facilities to computer programmers for software development.

Logistic regression - Logistic regression is a statistical model that in its basic form uses a logistic function to model a binary dependent variable, although many more complex extensions exist.

Linear regression - In statistics, linear regression is a linear approach to modelling the relationship between a scalar response and one or more explanatory variables (also known as dependent and independent variables).

Long Short-term Memory (LSTM) - is an artificial recurrent neural network (RNN) architecture used in the field of deep learning. Unlike standard feedforward neural networks, LSTM has feedback connections.

Keras - is an open-source software library that provides a Python interface for artificial neural networks. Keras acts as an interface for the TensorFlow library.

TensorFlow - is a free and open-source software library for machine learning

Time Series - Time series forecasting is used to predict future values based on previously observed values and one of the best tools for trend analysis and future prediction.

Survival Modeling - Survival analysis is a statistical method that aims to predict the time to an event, such as death, the diagnosis of a disease or the failure of a mechanical part.

Online Machine Learning - is a method of machine learning in which data becomes available in a sequential order and is used to update the best predictor for future data at each step, as opposed to batch learning techniques which generate the best predictor by learning on the entire training data set at once.

Neural Networks - are a series of algorithms that mimic the operations of a human brain to recognize relationships between vast amounts of data.