****

**CHAPTER 1**

**1.1 Introduction to Technologies**

**1.1.1 Angular 2**

Angular 2 is the new improved version of the ever popular JavaScript framework  AngularJs . Angular2 is a re-imagining of Angular applying all lessons learned from v1.x and promotes a component based architecture while leveraging new features of ES2015 (or TypeScript) like classes and modules.

* **Speed and Performance**

Achieve the maximum speed possible on the Web Platform today, and take it further, via Web Workers and server-side rendering.

* **Incredible Tooling**

Build features quickly with simple, declarative templates. Extend the template language with your own components and use a wide array of existing components. Get immediate Angular-specific help and feedback with nearly every IDE and editor. All this comes together so you can focus on building amazing apps rather than trying to make the code work.

* **Loved by Millions**

From prototype through global deployment, Angular delivers the productivity and scalable infrastructure that supports Google's largest applications.

**1.1.2 TypeScript**

**TypeScript** is a free and open source programming language developed and maintained by Microsoft. It is a strict superset of JavaScript, and adds optional static typing and class-based object-oriented programming to the language. Anders Hejlsberg , lead architect of  C#  and creator of  Delphi  and Turbo Pascal , has worked on the development of TypeScript .TypeScript may be used to develop JavaScript applications for client-side or server-side(Node.js) execution.

TypeScript is designed for development of large applications and transcompiles to JavaScript .As TypeScript is a superset of JavaScript, any existing JavaScript programs are also valid TypeScript programs.

**1.1.3 Node.js**

**Node.js** is an open-source, cross-platform JavaScript runtime environment for executing JavaScript code server-side, and uses the Chrome V8 JavaScript engine. Historically, JavaScript was used primarily for client-side scripting, in which scripts written in JavaScript are embedded in a webpage's HTML, to be run client-side by a JavaScript engine in the user's web browser. Node.js enables JavaScript to be used for server-side scripting, and runs scripts server-side to produce dynamic web page content before the page is sent to the user's web browser. Consequently, Node.js has become one of the foundational elements of the "JavaScript everywhere" paradigm ,allowing web application development to unify around a single programming language, rather than rely on a different language for writing server side scripts. Though .js is the conventional filename extension for JavaScript code, the name "Node.js" is not referring to a particular file in this context—it's just the name of the product.

**1.1.4 JavaScript**

**JavaScript** is a high-level , dynamic,untyped and interepted programming language .It has been standardized in the ECMAScript language specification . Alongside [HTML](https://en.wikipedia.org/wiki/HTML) and [CSS](https://en.wikipedia.org/wiki/CSS), JavaScript is one of the three core technologies of [World Wide Web](https://en.wikipedia.org/wiki/World_Wide_Web) [content production](https://en.wikipedia.org/wiki/Content_engineering); the majority of [websites](https://en.wikipedia.org/wiki/Website) employ it, and all modern [Web browsers](https://en.wikipedia.org/wiki/Web_browser) support it without the need for plug-ins.JavaScript is [prototype-based](https://en.wikipedia.org/wiki/Prototype-based_programming) with [first-class functions](https://en.wikipedia.org/wiki/First-class_function), making it a [multi-paradigm](https://en.wikipedia.org/wiki/Multi-paradigm) language,supporting [object-oriented](https://en.wikipedia.org/wiki/Object-oriented_programming), [imperative](https://en.wikipedia.org/wiki/Imperative_programming), and [functional](https://en.wikipedia.org/wiki/Functional_programming) [programming styles](https://en.wikipedia.org/wiki/Programming_paradigm). It has an [API](https://en.wikipedia.org/wiki/Application_programming_interface) for working with text, [arrays](https://en.wikipedia.org/wiki/Array_data_type), dates and [regular expressions](https://en.wikipedia.org/wiki/Regular_expression), but does not include any [I/O](https://en.wikipedia.org/wiki/Input/output), such as networking, storage, or graphics facilities, relying for these upon the host environment in which it is embedded.

Although there are strong outward similarities between JavaScript and Java, including language name, [syntax](https://en.wikipedia.org/wiki/Syntax_(programming_languages)), and respective [standard libraries](https://en.wikipedia.org/wiki/Standard_library), the two are distinct languages and differ greatly in their design. JavaScript was influenced by programming languages such as [Self](https://en.wikipedia.org/wiki/Self_(programming_language)) and [Scheme](https://en.wikipedia.org/wiki/Scheme_(programming_language)).

**1.1.5 Web Services**

A web service is a service offered by an electronic device to another electronic device, communicating with each other via the World Wide Web. In a Web service, Web technology such as HTTP, originally designed for human-to-machine communication, is utilized for machine-to-machine communication, more specifically for transferring machine readable file formats such as XML and JSON. In practice, the web service typically provides an object-oriented web-based interface to a database server, utilized for example by another web server, or by a mobile application, that provides a user interface to the end user. Another common application offered to the end user may be a mashup, where a web server consumes several web services at different machines, and compiles the content into one user interface.

Web services may use SOAP over HTTP protocol, allowing less costly interactions over the Internet than via proprietary solutions like EDI/B2B. Besides SOAP over HTTP, web services can also be implemented on other reliable transport mechanisms like FTP. In a 2002 document, the W3C Web Services Architecture Working Group defined a Web Services Architecture, requiring a standardized implementation of a "web service."

We can identify two major classes of web services:

* REST- compliant web services , in which the primary purpose of the service is to manipulate XML representations of web resources using a uniform set of "stateless" operations; and
* arbitrary web services , in which the service may expose an arbitrary set of operations.

— W3C, Web Services Architecture

**1.1.5.1 XML**

In computing, Extensible Markup Language (XML) is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable. The W3C's XML 1.0 Specification and several other related specifications—all of them free open standards—define XML.

The design goals of XML emphasize simplicity, generality, and usability across the Internet. It is a textual data format with strong support via Unicode for different human languages. Although the design of XML focuses on documents, the language is widely used for the representation of arbitrary data structures such as those used in web services.

**1.1.5.2 JSON**

In computing, JavaScript Object Notation or JSON , is an open-standard format that uses human-readable text to transmit data objects consisting of attribute–value pairs. It is the most common data format used for asynchronous browser/server communication, largely replacing XML, and is used by AJAX.

JSON is a language-independent data format. It was derived from JavaScript, but as of 2017 many programming languages include code to generate and parse JSON-format data. The official Internet media type for JSON is application/json. JSON filenames use the extension .json.

**1.1.5.2 HTTP**

The Hypertext Transfer Protocol (HTTP) is an application protocol for distributed, collaborative, and hypermedia information systems.[1] HTTP is the foundation of data communication for the World Wide Web.

Hypertext is structured text that uses logical links (hyperlinks) between nodes containing text. HTTP is the protocol to exchange or transfer hypertext.

**1.1.6 REST**

Representational state transfer (REST) or RESTful Web services are one way of providing interoperability between computer systems on the Internet. REST-compliant Web services allow requesting systems to access and manipulate textual representations of Web resources using a uniform and predefined set of stateless operations. Other forms of Web service exist, which expose their own arbitrary sets of operations such as WSDL and SOAP. "Web resources" were first defined on the World Wide Web as documents or files identified by their URLs, but today they have a much more generic and abstract definition encompassing every thing or entity that can be identified, named, addressed or handled, in any way whatsoever, on the Web. In a RESTful Web service, requests made to a resource's URI will elicit a response that may be in XML, HTML, JSON or some other defined format. The response may confirm that some alteration has been made to the stored resource, and it may provide hypertext links to other related resources or collections of resources. Using HTTP, as is most common, the kind of operations available include those predefined by the HTTP verbs GET, POST, PUT, DELETE and so on. By making use of a stateless protocol and standard operations, REST systems aim for fast performance, reliability, and the ability to grow, by re-using components that can be managed and updated without affecting the system as a whole, even while it is running.

**1.2 Introduction to Banking**

**1.2.1 Clearing House**

In banking and finance, clearing denotes all activities from the time a commitment is made for a transaction until it is settled. Clearing of payments is necessary to turn the promise of payment (for example, in the form of a cheque or electronic payment request) into the actual movement of money from one account to another.

In trading, clearing is necessary because the speed of trades is much faster than the cycle time for completing the underlying transaction. It involves the management of post-trading, pre-settlement credit exposures to ensure that trades are settled in accordance with market rules, even if a buyer or seller should become insolvent prior to settlement. Processes included in clearing are reporting/monitoring, risk margining, netting of trades to single positions, tax handling, and failure handling.

The Clearing House Association, L.L.C. is a New York-headquartered trade group and the nation’s first and oldest banking association representing 24 of the world's largest commercial banks, which collectively employ over 2 million people and hold more than half of all U.S. deposits. It is a nonpartisan organization that advocates on regulatory, legislative, and legal public policy issues on behalf of its Owner Banks before policymakers, courts of law, and standards setters in the U.S. and abroad.

The Clearing House seeks a level playing field among similarly situated market participants, in which a legal and regulatory framework promotes systemic stability, economic growth, and a safe and sound banking system. Unique among trades for its sole focus on large-scale commercial banking and payments issues, the Association and its Owner Banks form strong consensus positions on issues vital to the banking industry that are technically detailed and research- and data-driven.

**1.2.2 Pledge**

Pledge is used when the lender (pledgee) takes actual possession of assets i.e. certificates, goods ). Such securities or goods are movable securities. In this case the pledgee retains the possession of the goods until the pledgor (i.e. borrower) repays the entire debt amount. In case there is default by the borrower, the pledgee has a right to sell the goods in his possession and adjust its proceeds towards the amount due (i.e. principal and interest amount). Some examples of pledge are Gold /Jewellery Loans, Advance against goods,/stock, Advances against National Saving Certificates etc.

**1.2.2 Release**

A pledgor bank will only be allowed to release collateral when there are no longer any state funds on deposit or the current market value of any remaining collateral is equal to or greater than the maintenance percentage (110% of the amount deposited by the State Treasurer with the pledgor bank plus the interest due at maturity, in excess of the FDIC-insured limit).

**­­**

**CHAPTER 2**

**CHAPTER 3**

**3.Problem Definition and Background**

**3.1 Existing Approach**

DonorSnap provides you with tools to manage the pledge process from start to finish. A donation forecast report is available to help an organization try to estimate incoming receipts based upon initial pledge information. But the system is globally available and are not specific to the requirements of our client.

**3.2 Existing Approach**

In the proposed system, we create a standalone pledger management for our clients. It is easy to create a Pledge within the system and then to apply subsequent donation receipts to the pledge. The system allows you to monitor the status of a specific pledge or to view all pledge activity as a whole on printed reports.

The main purpose is to create a Pledge Management page which is used by administrators and staff members to manage pledge records. From this page, you can view clients who have pledged to a specific fund, add and delete pledge records, view pledge details, and make adjustments to existing pledges. The organization administrator can see all funds and their associated pledges while staff members can see only those pledges associated with the funds they granted permission to access. Every clients will be having their own broker dealer and every dealer will have their own unique IBD(Broker Dealer ID). Using that IBD, clients can either pledge or release shares and securities.

**3.3 Problem Statement**

A Stock Pledge is nothing but a transfer of stocks against a debt. It is an agreement between the client and the bank. The debtor pledges the shares as an asset against the amount of money taken from a lender and promises to return the amount within specific period. The debtor pledges the stocks as a security against the debt. According to the law, after the payment of the obligation the bank in which stocks are ledged must return the stocks to the debtor and the agreement stands void.

In Early, it was very difficult to manage this pledge process for the individual clients. To view the pledge details, the client has to navigate to different modules which are tedious process. Since, many components are needed to check the status of the individual user’s pledge process, the convenience of the user and the response time is affected.

In order to overcome all the burden of the user, a single page is developed to fetch all the details from various components and displayed in the required format. User can specify their own requirements and can change the columns to be displayed. Every shares and securities associated with each banks for the particular broker id can be viewed with great ease. The pledging process can be scalable for existing and new client.

**CHAPTER 4**

**4.Requirements**

**4.1. Hardware Support**

**4.1.1. Client Side:**

The most common set of requirements defined by any operating system or software application is the physical computer resources, also known as hardware. A hardware requirement list is often accompanied by a Hardware Compatibility List (HCL), especially in case of operating systems. An HCL lists tested, compatible and sometimes incompatible hardware devices for a particular operating system or application.

|  |  |
| --- | --- |
| **Processor** | Intel core 2 duo and advance |
| **Speed** | 2.0 GHz |
| **Hard Disk Drive** | 250 GB and above. |
| **Operating System** | Windows, linux |
| **Memory** | 2 GB RAM and above |
| **System Type** | 32,64 bit Operating System |

Table: 4.1.Hardware Specification

**4.1.2. Server Side:**

**Mainframe:**

A Mainframe Computer is a high performance Multi User computer system which is the most scalable, available, reliable and secured machine in the world capable of performing some Million Instructions per second (up to 569,632 MIPS).

**Characteristics:**

1) Reliable single-thread performance

2) Maximum I/O connectivity

3) Maximum I/O bandwidth

4) Reliability, Availability & Serviceability (RAS)

5) Unbreakable Security & Scalability (USS)

**Z13 Mainframe:**

The z13 processor has faster I/O and the ability to address up to 10144 GiB of RAIM memories -- three times as much as its predecessor. It can house up to 168 processor units in a single system and run as many as 8,000 virtual servers.

At a maximum 5GHz, the z13's processor is slower in terms of clock speed than the chip in the z12, but IBM says it more than compensates for that with other improvements. The chip has eight cores compared with six for its predecessor, and it's manufactured on a newer, 22 nanometer process, which should mean smaller, faster transistors.

The z13 supports up to 8000 Linux images simultaneously for cloud computing. For the mobile economy the z13 does real-time encryption and can process 2.5 billion transactions per day. For the z13, IBM spent over 1 billion dollars and five years of development and with more than 500 new patents.

|  |  |
| --- | --- |
| **Available** | March 9,2015 |
| **Memory** | Up to 10TB |
| **Number Of Models** | 5-NE1,NC9,N96,N63,N30 |
| **Channels** | -PCIe Gen3 16GBps channel buses  -SIX CSSs, upto 85 LPARs  -4 Sub channels sets per CSS  -Flash Express |
| **Operating Systems** | z/OS, z/VM, z/VSE, z/TPF, Linux on z Systems |

Table:4.2. Mainframe Specification

**4.2. Software Requirements**

The list of software required for the project is:

(i) Windows Operating System

(ii) Java8

(iii)Weka 3.8

(iv)R Studio

**4.2.1. Windows Operating System**

The project was decided to be developed using Java and R language. JVM is available for Linux, Windows, MAC, Solaris operating systems. Out of those, we decided to choose Windows 8 operating system,, as it is more stable and reliable with long term support.

**4.2.2. Java8**

Java is platform independent meaning project compiled in one machine can be run in any other machine that has a version of JVM available for it regardless of the underlying architecture. It is an object oriented GPL licensed programming language.

Java 8 is the only version of Java that is currently supported. Main reason for choosing Java 8 is because of ease of prototyping in it.

**4.2.3 Weka 3.8**

Weka is a collection of machine learning algorithms for data mining tasks. The algorithms can either be applied directly to a dataset or called from your own Java code. Weka contains tools for data pre-processing, classification, regression, clustering, association rules, and visualization. It is also well-suited for developing new machine learning schemes.Weka is a workbench that contains a collection of visualization tools and algorithms for data analysis and predictive modeling, together with graphical user interfaces for easy access to these functions. This original version was primarily designed as a tool for analyzing data from agricultural domains,but the more recent fully Java-based version (Weka 3.8), for which development started in 1997, is now used in many different application areas, in particular for educational purposes and research.

**Advantages of Weka include:**

* Free availability under the GNU General Public License.
* Portability, since it is fully implemented in the Java programming language and thus runs on almost any modern computing platform.
* A comprehensive collection of data preprocessing and modeling techniques.
* Ease of use due to its graphical user interfaces.

**Weka packages:**

|  |  |  |
| --- | --- | --- |
| **Package Name** | **Purpose** | **Usage** |
| Libsvm | Classification, Regression | A wrapper class for the libsvm tools |
| WekaExcel | Converter | WEKA MS Excel loader/saver |
| timeSeriesFilters | Filters, Time Series | Time Series Filters |
| timeseriesForecasting | Time series | Time series forecasting environment. |

**4.2.4 R Studio**

R is a programming language and software environment for statistical computing and graphics supported by the R Foundation for Statistical Computing. The R language is widely used among statisticians and data miners for developing statistical software and data analysis. Polls, surveys of data miners, and studies of scholarly literature databases show that R's popularity has increased substantially in recent years.

**Package:**e1071

**Package description:** Functions for latent class analysis, short time Fourier transform, fuzzy clustering, support vector machines, shortest path computation, bagged clustering, naive Bayes classifier

**4.3. Dataset**

The dataset for our experiment were crawled from UCI repository.

*Table 1:Stock market dataset*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Date | open | high | low | close |
| 01-Jan-14 | 10312.81 | 10338.24 | 10254.12 | 10269.29 |
| 02-Jan-14 | 10272.36 | 10365.12 | 9936.96 | 9977.82 |
| 03-Jan-14 | 9900.59 | 9936.77 | 9760.24 | 9807.16 |
| 06-Jan-14 | 9812.43 | 9912.25 | 9752.26 | 9820.78 |
| 07-Jan-14 | 9855.68 | 9970.51 | 9803.99 | 9869.15 |
| 08-Jan-14 | 9910.21 | 9924.71 | 9715.71 | 9735.82 |
| 09-Jan-14 | 9750.72 | 9795.46 | 9525.58 | 9537.91 |
| 10-Jan-14 | 9548.05 | 9640.07 | 9392.79 | 9415.33 |
| 13-Jan-14 | 9443.98 | 9595.85 | 9443.98 | 9561.01 |
| 14-Jan-14 | 9568.6 | 9686.44 | 9469.67 | 9501.12 |
| 15-Jan-14 | 9546.92 | 0 | 9542.86 | 9689.03 |
| 16-Jan-14 | 9748.37 | 9798.14 | 9655.4 | 9713.12 |
| 17-Jan-14 | 9702.9 | 9758.83 | 9656.47 | 9688.56 |
| 20-Jan-14 | 9695.81 | 9765.87 | 9632.31 | 9700.22 |
| 21-Jan-14 | 9745.99 | 9820.37 | 9743.63 | 9761.91 |
| 22-Jan-14 | 9759.99 | 9803.19 | 9676.27 | 9710.82 |
| 23-Jan-14 | 9808.28 | 9969.64 | 9808.28 | 0 |
| 24-Jan-14 | 9836.14 | 9857.61 | 9609.81 | 9624.23 |

**CHAPTER 5**

**5.Proposed Approach**

**5.1 Overview of Proposed Design**

Normalization is accomplished through applying some formal rules either by a process of synthesis or decomposition. a non-linear function is leaned by linear learning machine mapping into high dimensional kernel induced feature space.Regression is that presenting the solution by means of small subset of training points.The below figure 3 gives the overview of proposed design.

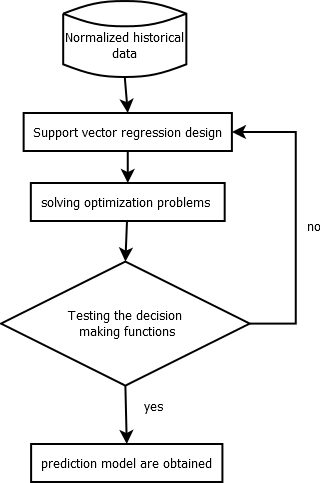


Figure 3: Proposed Design

**5.2 Normalizing historical data**

Database normalization, or simply normalization, is the process of organizing the [columns](https://en.wikipedia.org/wiki/Column_(database)) (attributes) and [tables](https://en.wikipedia.org/wiki/Table_(database)) (relations) of a [relational database](https://en.wikipedia.org/wiki/Relational_database) to reduce [data redundancy](https://en.wikipedia.org/wiki/Data_redundancy)and improve data integrity.

Normalization involves arranging attributes in tables based on [dependencies](https://en.wikipedia.org/wiki/Dependency_theory_(database_theory)) between attributes, ensuring that the dependencies are properly enforced by database integrity constraints. Normalization is accomplished through applying some formal rules either by a process of synthesis or decomposition. Synthesis creates a normalized database design based on a known set of dependencies. Decomposition takes an existing (insufficiently normalized) database design and improves it based on the known set of dependencies.

When a fully normalized database structure is extended to allow it to accommodate new types of data, the pre-existing aspects of the database structure can remain largely or entirely unchanged. As a result, applications interacting with the database are minimally affected.

**5.3 Support Vector Regression Design**

Support Vector Machine can be applied not only to classification problems but also to the case of regression. Still it contains all the main features that characterize maximum margin algorithm: a non-linear function is leaned by linear learning machine mapping into high dimensional kernel induced feature space. The capacity of the system is controlled by parameters that do not depend on the dimensionality of feature space.

In the same way as with classification approach there is motivation to seek and optimize the generalization bounds given for regression. They relied on defining the loss function that ignores errors, which are situated within the certain distance of the true value. This type of function is often called – epsilon intensive – loss function.

**5.4 Solving Optimization Problems**

One of the most important ideas in Support Vector Classification and Regression cases is that presenting the solution by means of small subset of training points gives enormous computational advantages. Using the epsilon intensive loss function we ensure existence of the global minimum and at the same time optimization of reliable generalization bound.

In SVM regression, the inputIMG_256is first mapped onto a *m*-dimensional feature space using some fixed (nonlinear) mapping, and then a linear model is constructed in this feature space.Using mathematical notation, the linear model (in the feature space) *f(x,w)* is given by



Where gj(x) ,j=1,2,…m denotes a set of nonlinear transformations

*b* is the “bias” term

SVM regression performs linear regression in the high-dimension feature space using ξ-insensitive loss and, at the same time, tries to reduce model complexity by minimizing *w2*. This can be described by introducing (non-negative) slack variables ξi , i=1…n,to measure the deviation of training samples outside *w*-insensitive zone.

Thus SVM regression is formulated as minimization of the following functional:



**5.5Algorithm Description**

**5.5.1 Algorithm 1 -Support Vector Regression**

In our SVR algorithm we have given a set of input vectors and associated responses, and it fits a model to try and predict the response given a new input vector. Consider we have stock prices for N days. Then, for each day we could construct a feature vector, which, in a simple case, could be be the previous day's price and the current day's price. The response for each feature vector would be the next day's price. Thus, given yesterday's price and today's price the objective would be to predict the next days price.With a time series, an import step is determining what your "feature vector" x will be; each x is called a "feature" and can be calculated from present or past data, and each y, the response, will be the future change over some time period of whatever you're trying to predict.

Below is the code to make predictions with Support Vector Regression:

* reg.model <- lm(open~Date,data=t1)
* t1$forecast.open.predicted <- predict(reg.model,data=t1)
* points(t1$Date, t1$forecast.open.predicted,

col = "green",

pch=18)

The function will automatically choose SVM if it detects that the data is categorical.

This time the predictions is closer to the real values.

* error <- data$open - predictedY
* svrPredictionRMSE <- rmse(error)

As expected the RMSE is better than the tuned Support vector regression values.

**5.5.2 Algorithm 2 – Simple Moving Average**

In financial applications a simple moving average (SMA) is the unweighted [mean](https://en.wikipedia.org/wiki/Arithmetic_mean) of the previous *n* data. However, in science and engineering the mean is normally taken from an equal number of data on either side of a central value. This ensures that variations in the mean are aligned with the variations in the data rather than being shifted in time. An example of a simple equally weighted running mean for a n-day sample of closing price is the mean of the previous *n* days' closing prices. If those prices are *pm,pm-1,…..pm-(n-1)*





The period selected depends on the type of movement of interest, such as short, intermediate, or long-term. In financial terms moving-average levels can be interpreted as [support](https://en.wikipedia.org/wiki/Support_(technical_analysis)) in a falling market, or [resistance](https://en.wikipedia.org/wiki/Resistance_(technical_analysis)) in a rising market.

If the data used are not centered around the mean, a simple moving average lags behind the latest datum point by half the sample width. An SMA can also be disproportionately influenced by old datum points dropping out or new data coming in. One characteristic of the SMA is that if the data have a periodic fluctuation, then applying an SMA of that period will eliminate that variation

**CHAPTER 6**

**6. Implement­ation**

**6.1 Weka**

**6.1.1 Preprocessing**

The dataset used here is the stock market data which available in arff format.

* Loading the data. We can load the dataset into weka by clicking on open button in preprocessing interface and selecting the appropriate file.
* Once the data is loaded, weka will recognize the attributes and during the scan of the data weka will compute some basic strategies on each attribute. The left panel shows the list of recognized attributes while the top panel indicates the names of the base relation or table and the current working relation .
* Clicking on an attribute in the left panel will show the basic statistics on the attributes for the categorical attributes the frequency of each attribute value is shown, while for continuous attributes we can obtain min, max, mean, standard deviation and deviation etc.,
* The visualization in the right button panel in the form of cross-tabulation across two attributes. Note:we can select another attribute using the dropdown list.
* Selecting or filtering attributes Removing an attribute-When we need to remove an attribute,we can do this by using the attribute filters in weka.In the filter model panel,click on choose button,This will show a popup window with a list of available filters. Scroll down the list and select the “weka.filters.unsupervised.attribute.remove” filters.
* Next click the textbox immediately to the right of the choose button.In the resulting dialog box enter the index of the attribute to be filtered out.
* Make sure that invert selection option is set to false.The click OK now in the filter box.you will see “Remove-R-7”.
* Click the apply button to apply filter to this data.This will remove the attribute and create new working relation.
* Save the new working relation as an arff file by clicking save button on the top panel.(stockmarket.arff)

*Table 2:Preprocessed dataset*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Date | open | high | low | close |
| 01-Jan-14 | 10312.81 | 10338.24 | 10254.12 | 10269.29 |
| 02-Jan-14 | 10272.36 | 10365.12 | 9936.96 | 9977.82 |
| 03-Jan-14 | 9900.59 | 9936.77 | 9760.24 | 9807.16 |
| 06-Jan-14 | 9812.43 | 9912.25 | 9752.26 | 9820.78 |
| 07-Jan-14 | 9855.68 | 9970.51 | 9803.99 | 9869.15 |
| 08-Jan-14 | 9910.21 | 9924.71 | 9715.71 | 9735.82 |
| 09-Jan-14 | 9750.72 | 9795.46 | 9525.58 | 9537.91 |
| 10-Jan-14 | 9548.05 | 9640.07 | 9392.79 | 9415.33 |
| 13-Jan-14 | 9443.98 | 9595.85 | 9443.98 | 9561.01 |
| 14-Jan-14 | 9568.6 | 9686.44 | 9469.67 | 9501.12 |
| 15-Jan-14 | 9546.92 | 9703.9 | 9542.86 | 9689.03 |
| 16-Jan-14 | 9748.37 | 9798.14 | 9655.4 | 9713.12 |
| 17-Jan-14 | 9702.9 | 9758.83 | 9656.47 | 9688.56 |
| 20-Jan-14 | 9695.81 | 9765.87 | 9632.31 | 9700.22 |
| 21-Jan-14 | 9745.99 | 9820.37 | 9743.63 | 9761.91 |
| 22-Jan-14 | 9759.99 | 9803.19 | 9676.27 | 9710.82 |
| 23-Jan-14 | 9808.28 | 9969.64 | 9808.28 | 9889.53 |
| 24-Jan-14 | 9836.14 | 9857.61 | 9609.81 | 9624.23 |

**6.1.2. SMOreg**

SMOreg is a function used for forecasting the result in regression.It implements the support vector machine for regression.

**Class column**

Choose the column that contains the target variable.

**Preliminary Attribute Check**

The Preliminary Attribute Check tests the underlying classifier against the DataTable specification at the inport of the node. Columns that are compatible with the classifier are marked with a green 'ok'. Columns which are potentially not compatible are assigned a red error message.

**Classifier Options**

C: The complexity constant C. (default 1)

N:Whether to 0=normalize/1=standardize/2=neither. (default 0=normalize)

I:Optimizer class used for solving quadratic optimization problem (default weka.classifiers.functions.supportVector.RegSMOImproved)

K:The Kernel to use.

(default: weka.classifiers.functions.supportVector.PolyKernel)

T: The tolerance parameter for checking the stopping criterion.

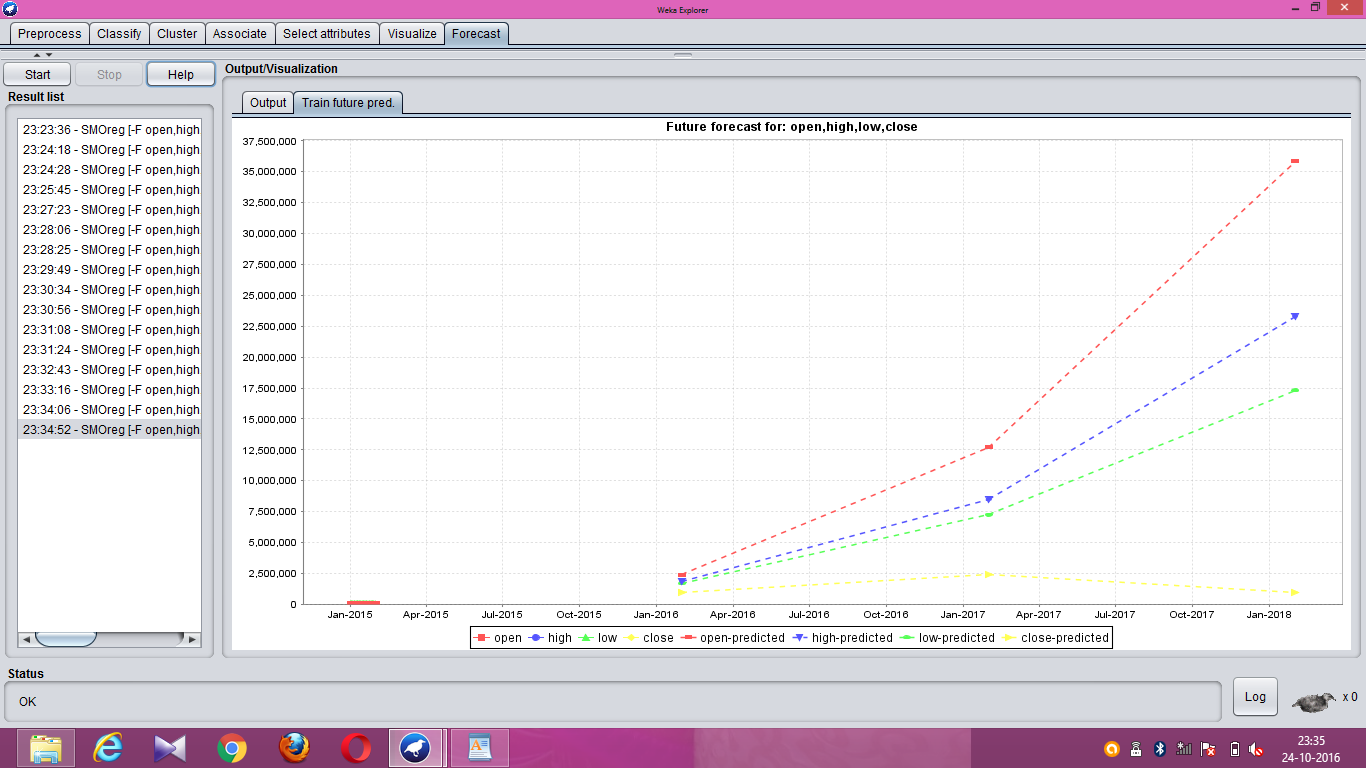
(default 0.001)

P: The epsilon for round-off error. (default 1.0e-12)

L: The epsilon parameter in epsilon-insensitive loss function.

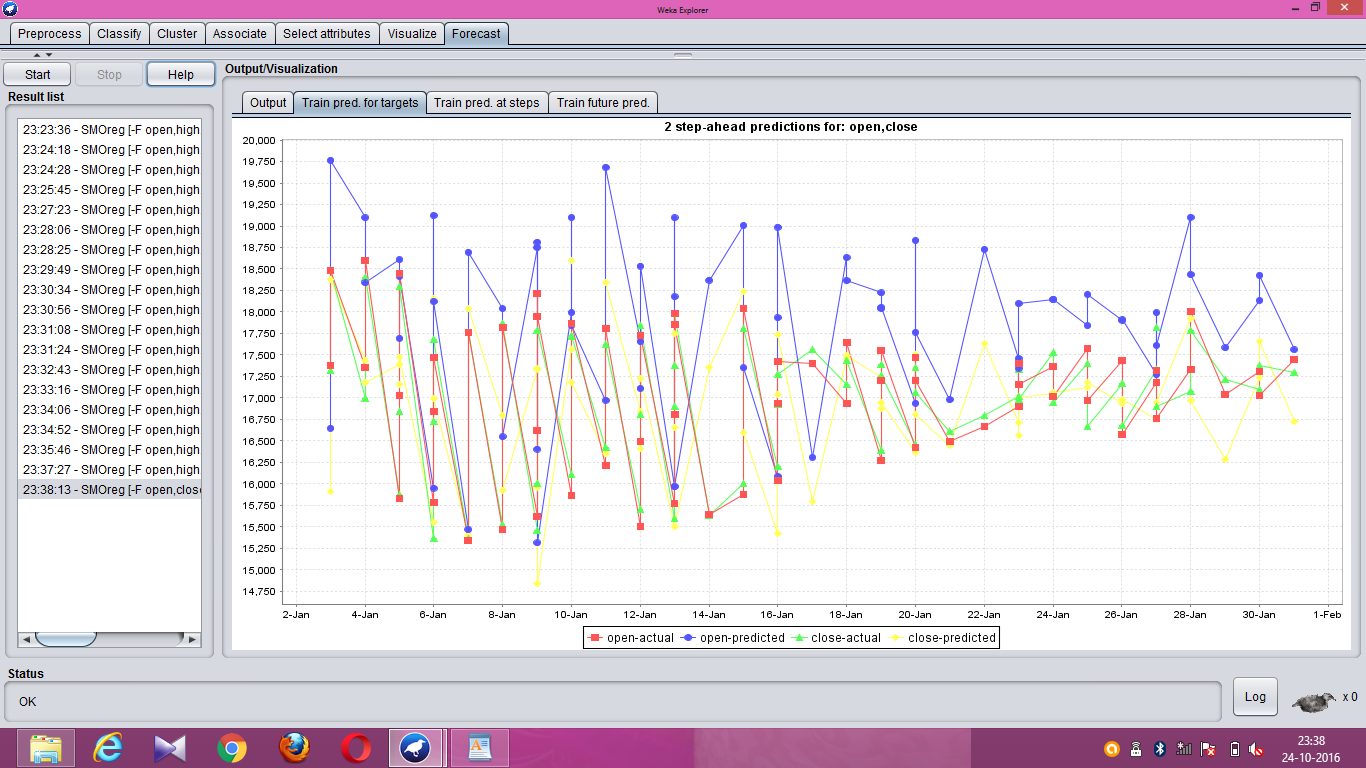
(default 1.0e-3).

The future forecast for all the attributes in stock market data is forecasted using weka and result is show below:

**

*Figure 4 :Future forecast using stock market data*

The future forecast for open and close attributes in stock market data is forecasted using weka and result is show below:

**

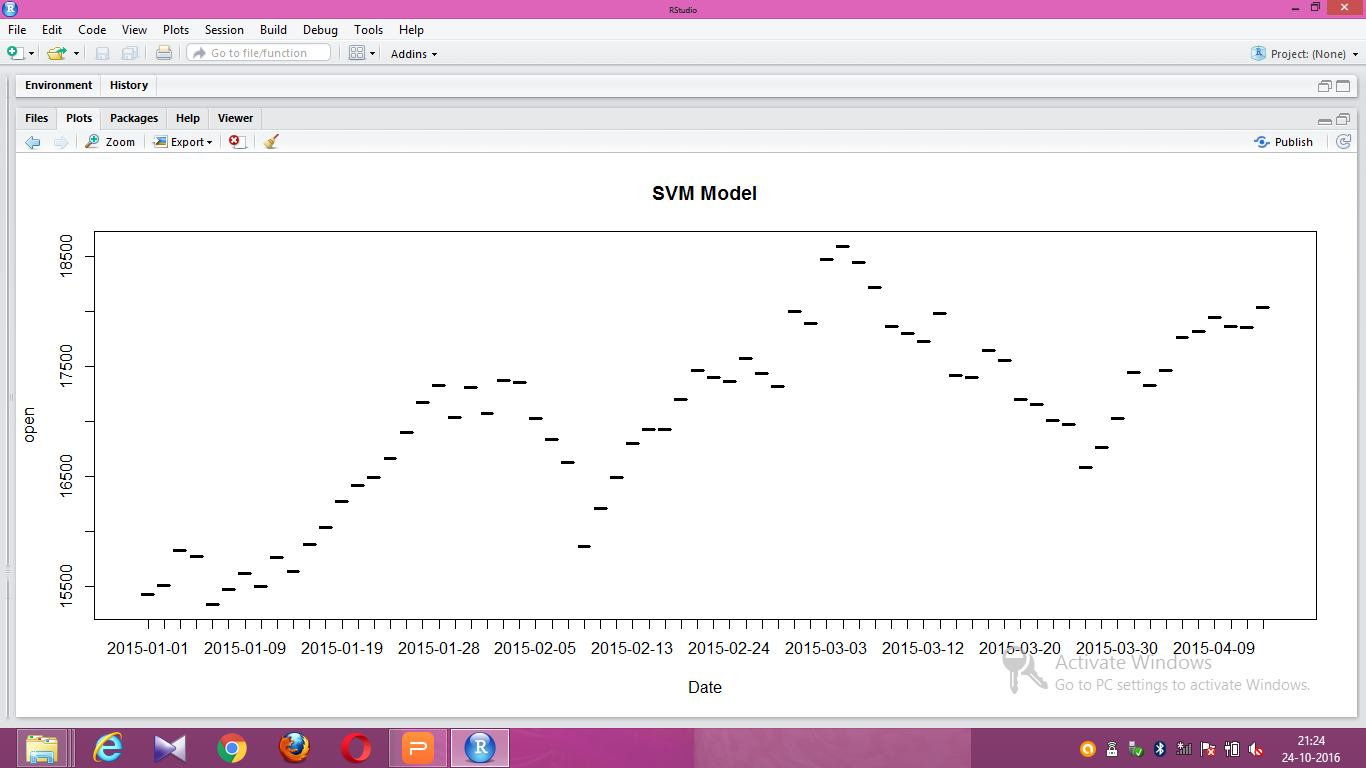
*Figure 5 Prediction of open and close attribute in stock market data*

**6.2 Support vector machine**

**Proposed procedure for SVM:**

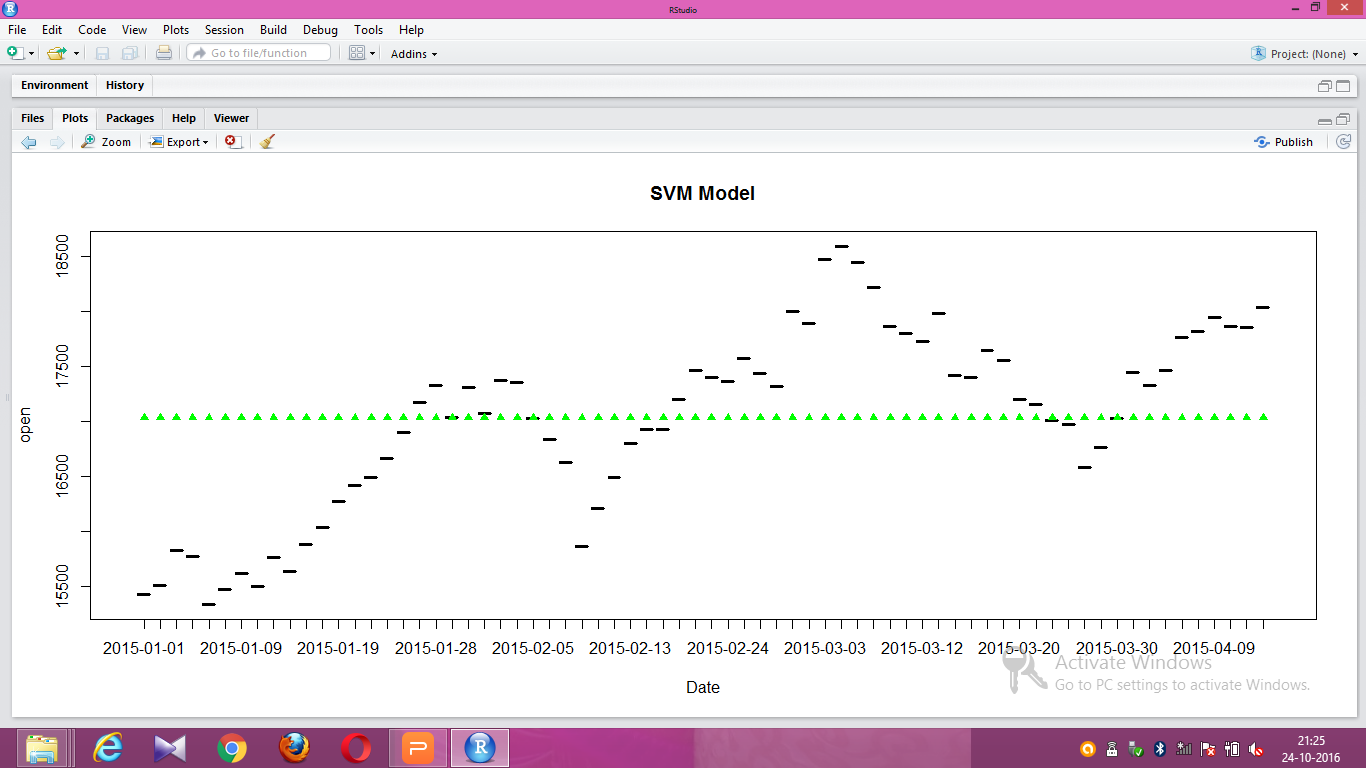
* Transform data to the format of an SVM package
* Conduct simple scaling on the data
* Consider the RBF kernel K(x, y) = e −γkx−yk
* Use cross-validation to find the best parameter C and γ
* Use the best parameter C and γ to train the whole training set
* Test

The future forecast for all the attributes in stock market data is forecasted using svm and result is show below:



*Figure 6:SVM model experimented in R*

The future forecast for open attribute in stock market data is forecasted using svm and result is show below:



*Figure 7:Prediction of open attribute in stock market data using SVM*

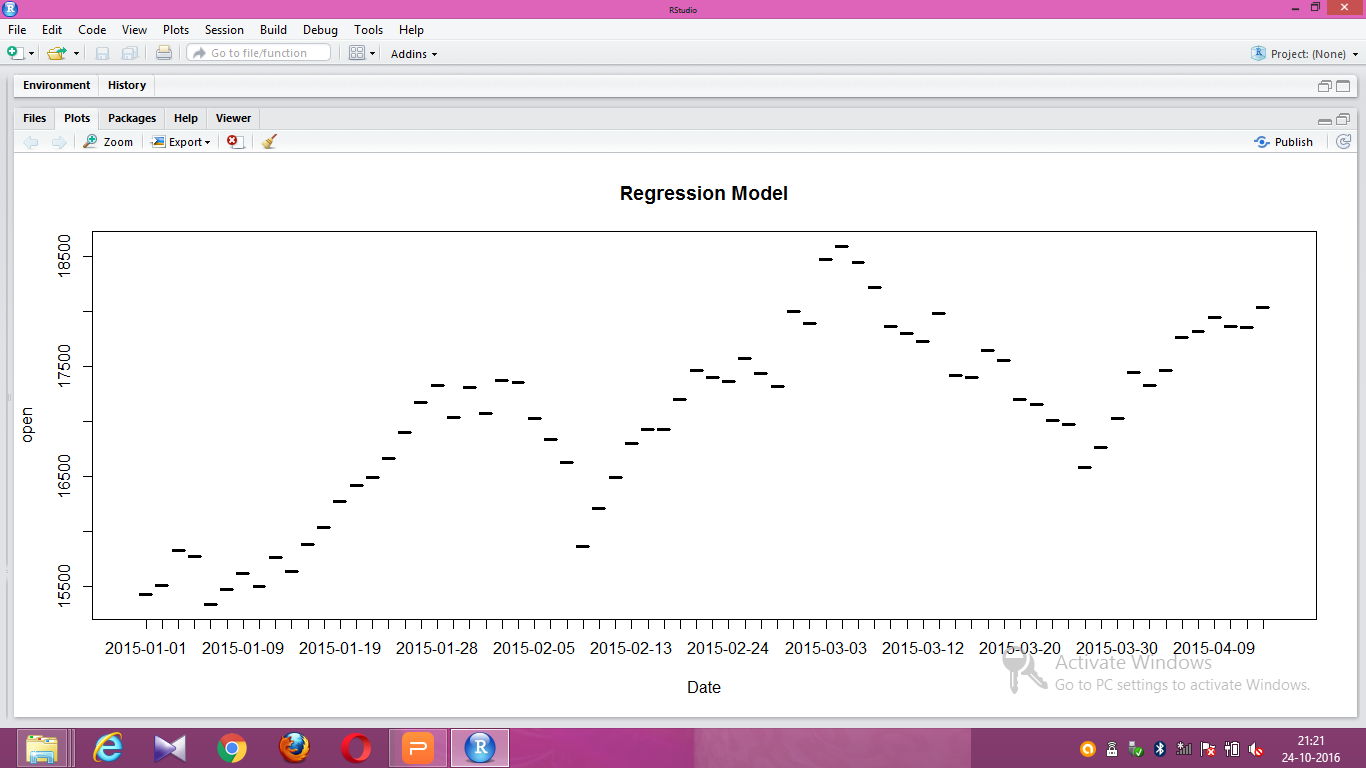
* Since,the selection of the kernel function parameters - for Gaussian kernels the width parameter [sigma] - and the value of [epsilon] in the [epsilon]-insensitive loss function.
* However, from a practical point of view perhaps the most serious problem with SVMs is the high algorithmic complexity and extensive memory requirements of the required quadratic programming in large-scale tasks.

**CHAPTER 7**

**7 Experiments and Results**

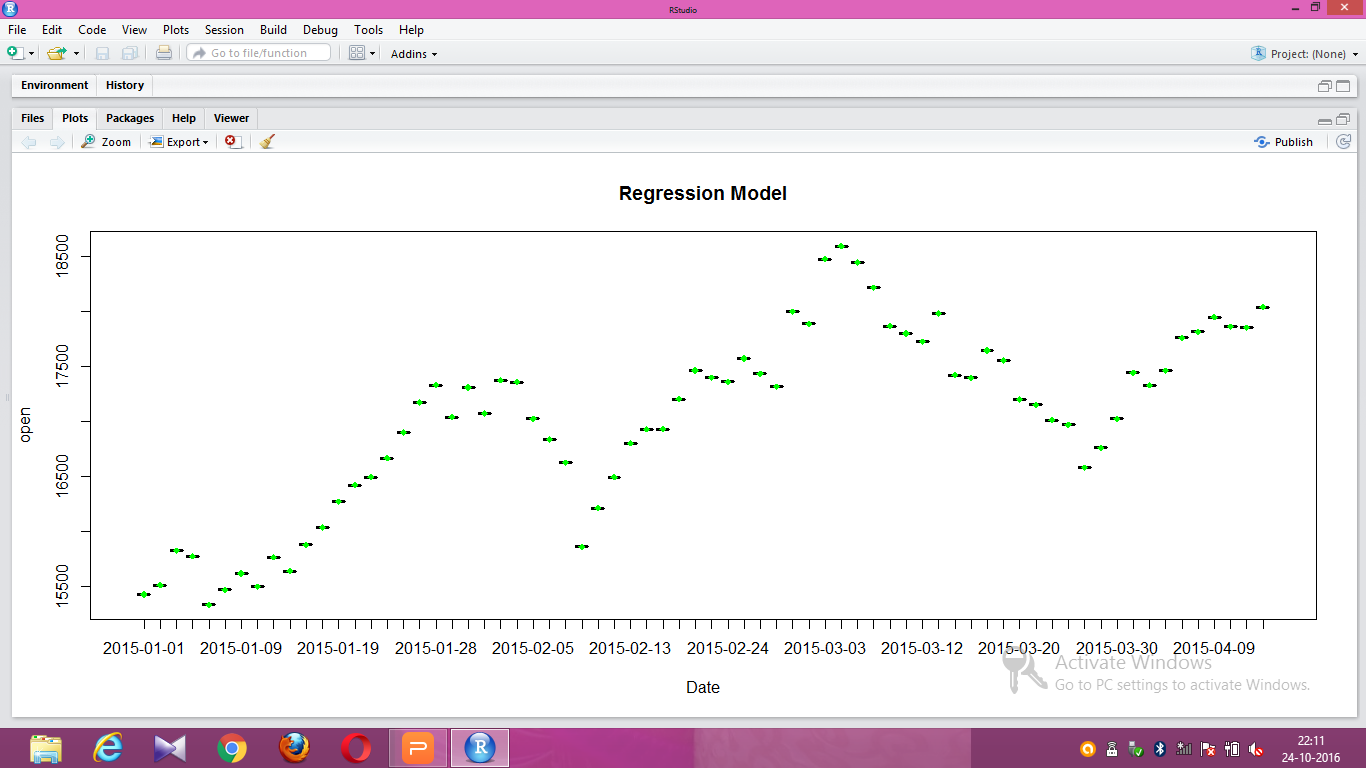
**7.1 Experiments**

**7.1.1 Regression model**

Support Vector Regression attempts to minimize the generalization error bound so as to achieve generalized performance. The idea of SVR is based on the computation of a linear regression function.The future forecast for all the attributes in stock market data is forecasted using svr and result is show below:

*Figure 8: Regression model of stock market data experimented in R*

The future forecast for open attribute in stock market data is forecasted using svm and result is show below:

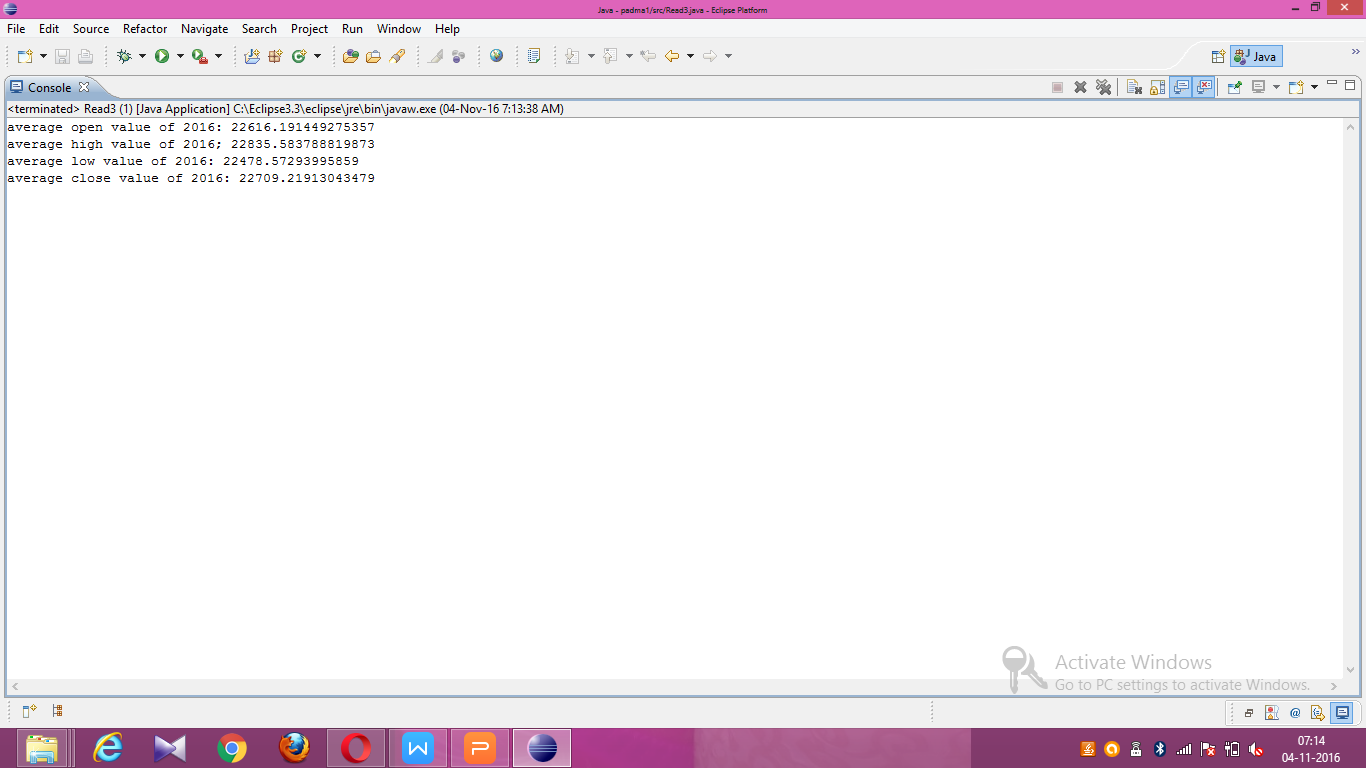


*Figure 9: Prediction of open attribute in stock market data using SVR*

**7.1.2 Moving Average Method**

Moving average is a calculation to analyze data points by creating series of averages of different subsets of the full data set.

The future forecast for all the attributes in stock market data is forecasted using moving average method and result is show below:

**

*Figure 10: Moving Average predicted value*

**7.2 Results Analysis**

The predicted value obtained from the Support vector regression and Moving Average method gives appropriate values of the present years and the future values is also predicted for the stock market.Our method was able to predict the shares around two years with a very high accuracy. Our methodology provided an average accuracy of 87%.





*Figure 11: Comparison of RMSE in MA and SVR*

**CHAPTER 8**

**8. Conclusion**

In this paper we predicted the future values in the stock market using Support vector regression and Moving Average method.The predicted value obtained from the algorithm will be nearly equal to the future values in the stock market.

**CHAPTER 9**

**9. REFERENCES**

1. *Albert Y. Chen,Tsung-Yu Lu,”Demand Forecast Using Data Analytics for the Preallocation of Ambulances”,.IEEE Journal of Biomedical and Health Infomatics, vol. 20, no. 4, July 2016.*
2. *Aditya N. Deshmukh, M. C. Deo, Prasad K. Bhaskaran, T. M. Balakrishnan Nair, and K. G. Sandhya,”Neural-Network-Based Data Assimilation to Improve Numerical Ocean Wave Forecast,IEEE Journal of Oceanic Engineering”, vol. 41, no. 4, October 2016.*
3. *Pritam Bose, Nikola K. Kasabov, Lorenzo Bruzzone, Fellow,”Spiking Neural Networks for Crop Yield Estimation Based on Spatio temporal Analysis of Image Time Series”,IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 11, November 2016.*
4. *Ye Ren, Ponnuthurai Nagaratnam Suganthan, and Narasimalu Srikanth,”A Novel Empirical Mode Decomposition With Support Vector Regression for Wind Speed Forecasting",IEEE Transactions on Neural Networks and Learning Systems, vol. 27, no. 8, August 2016.*
5. *Amin Hedayati Moghaddama, Moein Hedayati Moghaddamb, Morteza Esfandyaric.”Stock market index prediction using artificial neural network”,Journal of Economics, Finance and Administrative Science 21 (2016) 89–93.*