Advances in Mathematical & Computational Sciences Journal,
Vol. 5 No. 1,, March, 2017

Article Citation Format
Osang, F.B & Umoren, I. (2017): Performance Modelling and Optimization of Business Investment Using Decision Tree Approach.
Journal of Advances in Mathematical & Computational Sciences Vol. 5, No. 1.
Pp 1-24

Article Progression Time Stamps

Article Type: Research Article
Manuscript Received 17th February, 2017
Final Acceptance: 12th March, 2017
Article DOI: dx.doi.org/10.22624/AIMS/MATHS/V5N1P1

Performance Modelling and Optimization of Business Investment Using Decision Tree Approach

[1]Osang, Francis Bukie & [2]Umoren, Imeh
[1]Department of Computer Science, National Open University of Nigeria, Abuja, Nigeria
[2] Department of Computer Science, Akwa Ibom State University, Mkpat Enin PMB 1167, Akwa Ibom State, Nigeria
E-mails: [1]fosang@noun.edu.ng; [2]imehumoren@aksu.edu.ng
Phone: +234 705 7129 566
[1]Corresponding Author

ABSTRACT

Various optimization models used for business investments are considered to be complex in interpretation and required high level of expertise to understand. This poised a major setback for non-experienced and non-professional investors which make up the bulk of investors worldwide. The Decision Tree Model (DTM) is adopted to solve these problems using Decision Analysis Technique. The Decision Analysis Technique is implemented using a Decision Tree Model (DTM). The overall objective of this research is to optimize business investments by developing a trouble-free business investment optimization scheme. We investigate business investment and observed that the Decision Tree Model (DTM) offers possible optimization model to read, realize and interpret due to its graphical presentation of data, variables, and results. The field data was obtained from secondary sources consisting mainly of data from stock market monitoring. The data consist of twenty stock investments drawn from five different business sectors in Nigeria. Analysis was carried out on the data and subjected to Decisions Tree Model based on the weight of the Expected Monetary Value **(EMV)** of each investment. The results revealed that; investment with the highest payoff does not always maximize profit, the investment with the highest cash flow does not always maximize profit and investment that maximizes profit does not always carry the lowest weight of possible financial loss in all cases. Hence, we proposed that investors should use the Decision Tree Model (DTM) to optimize their business investment, adopt the Value Investing Principle (VIP) to determine the capital intensity of a potential investment. Nevertheless, investors who are averse to risk, should apply the maximum decision rule to their investment analysis. The scheme simplifies the tedious process of business investment optimization for various classes of investors and adjudge essential for safely introducing novices and non-professional investors to gain expertise of business investment optimization goals.

Keywords: Performance Modelling, Optimization, Business Investment, Decision Tree Approach.

1. INTRODUCTION

In recent times, there has been serious economic meltdown and recession in the world, corporate organizations, private investors and individuals are faced with the challenge of making intelligent business investment decisions that maximize their return on investment (ROI) and minimize all the risks associated with their investments. Investors are saddled with the responsibility of choosing investments that will optimize their portfolio. Making investment decisions in a world that has been hit by series of economic disasters, unforeseen circumstances such as sanctions, political unrest, etc. can be a daunting and difficult task for most investors. Various investment principles have been advocated by different individuals in the past but it has been observed that these principles are not tools for investment Decision Analysis; instead they are a set of guidelines which influence an investor's investment approach.

Investment principles alone are not sufficient to proffer intelligent business investment decisions as they lack certain aspects of information in a business that are paramount in investment decisions and this generates a problem of data quality. Nevertheless the quality of investment decision made by an investor is dependent on data quality. The quality of data acquired and analyzed by an investor must be consistent, valid, complete, accurate and uniform. Data quality has been a major challenge for investors; especially corporate investors who have turned to business intelligent systems to aid them make intelligent business investment decisions.

Many investors strive to have an optimized set of business investments(mostly referred to as portfolio optimization) using different portfolio optimization models and yet they are still looking for a portfolio optimization model that simplifies the optimization process and still captures real-life investment scenarios at the same time. The field of Operations Research has been exploited by investors in an attempt to improve investment decision making. Techniques such as optimization (or Mathematical Optimization), data envelopment analysis, econometric methods, analytical hierarchy process, Decision Analysis, etc. have been used in Operations Research to reach optimal solution to complex decision making problems of investment selection. Investment decision making is a problem of choosing between alternatives as a result of complexity in choice of investment and it is imperative that Decision Analysis is used as a tool to analyze investments and make decisions.

Decision Analysis approach comprises of principle, theory, methodology, and professional skills and practice which are used to address important decisions in a formal manner. Decision Analysis problem commonly utilises graphical tools such as influence diagrams and decision trees to present alternatives available to investors, the uncertainty they face, and evaluation measures that represent how well they achieve or optimize their investment objectives. Mathematical Optimization has been used by investors to determine whether to invest in a certain business or not. But Mathematical Optimization does not efficiently and effectively analyse investment problems with multiple variables and are very complex and difficult to understand by non-professional business investors. The focus of this research is to adopt Decision Analysis Methodology using the Decision Tree Model (DTM) to help investors analyse the multiple variables that influences investment decisions and make choices between competing investment alternatives that will optimize their investment objectives and achieve their investment goals.

1.1 Statement of Problem

Mathematical Optimization has been the most used technique in the areas of investment and portfolio optimization. Considering the complexity of the Mathematical Optimization technique and its difficulty of use and understanding. Non-professional investors who make up the bulk of investors' population have resorted to the application of inefficient

techniques such as predicting, hunches, and emotional disposition or biases to investment selection, which they fail to optimize their investments efficiently. Hence, this research paper sought to apply the Decision Analysis Technique to investment optimization in an attempt to provide investors with an optimization model that is easy to use, understand, interpret and manipulate, while leaving investors with an optimized solution to the business investment selection problems. Given that investors need to make efficient investment decisions in the face of enormous investment alternatives, the need to consider Decision Analysis Technique **(Decision Tree Model),** which can be used to analyze and weigh these competing investment alternatives.

2. REVIEW OF LITERATURE

When current income exceeds current consumption desires, people tend to save the excess. They can do several things with the savings. One possibility is to put the money under a mattress or bury it in the backyard until some future time when consumption desires exceed current income. When they retrieve their savings from the mattress or backyard, they have the same amount they saved. Another possibility is that they can give up the immediate possession of these savings for a future larger amount of money that will be available for future consumption. The tradeoff of present consumption for a higher level of future consumption is the reason for saving. What is done with the savings to make them increase over time is investment (Reilly & Brown, 2002). Investment is the purchase of an asset or item with the hope that it will generate income or appreciate in the future and be sold at a higher price (Wikipedia, Investment, 2015). Investment is the placement of capital in expectation of deriving income or profit from its use (Wiktionary, 2015). Markowitz (1952) introduced the Mean Variance (MV) model based on the principle that a portfolio is optimal if it maximizes the return for a given level of risk.

Konno and Yamakazi (1991) proposed a linear programming model which used the Mean-Absolute Deviation (MAD) as the risk function to replace the risk function in Markowitz's model. The model assumes there is no particular distribution for asset returns and is equivalent to the Markowitz model when they have a multivariate normal distribution. Researchers constructed a cardinality-constrained Markowitz model incorporating floor and ceiling constraints which was solved using genetic algorithms, simulated annealing, and tabu search. Costs were not addressed in this model (Chang, Meade, Beasley & Sharaiha, 1998). Busetti (2000) found that it is often too difficult to incorporate real-world constraints and dilemma into the classical theory (Markowitz's Mean Variance model), and that the weakness of the model is its assumption of multivariate normality. Gilli and Këllezi (2000) stated that another approach to portfolio selection optimization is scenario generation. This approach is a scenario analysis where uncertainty about future returns is modeled through a set of possible realizations, called scenarios. Scenarios of future outcome can be generated lying on a model, past returns or experts' advice.

Kellerer et al (2000) used mixed integer programming to construct heuristics for handling fixed costs. In their model they allow linear objective function and linear constraints on the transaction amounts. Consequently, non-linear constraints such as those on covariance and short fall risk cannot be handled in this model. Gilli (2004) stated that the rationale for the Mean Varian model introduced by Markowitz is to find the most attractive combination of return and risk where return is the mean of future expected returns and risk is the variance of the future gains. Papahristodoulou and Dotzauer (2004) compared the Linear Mean-absolute Deviation (LMAD) model and the mean variance (MV) model with out-of-sample data from shares traded in the Stockholm stock Exchange. They discovered that the MV model yields higher utility levels and higher degrees of risk aversion in very similar computing times. In a research carried out by Arnott and Wagner in 1990, they found that ignoring transaction cost results in inefficient portfolio, and this conclusion has been supported by Yoshimoto's 1996 research on empirical analysis (Fang, Lai & Wang, 2008).

The arguments of Koochakzadeh (2013) on the nature of the portfolio selection problem proposed several approaches as the stock portfolio optimization process involves forecasting the performance and the volatility of stocks as well as models for using these predictions in order to obtain a portfolio that suits the investor's preference profile. A solution for stock portfolio optimization involves building models from past available data to forecast the performance and the volatility of assets in future, in order to obtain the optimum portfolio for an individual's preferences. According to Mansourfar (2013), Modern Portfolio Theory suggests that investors should not select their assets merely according to the individual characteristics.

Nevertheless this research adopts the Decision Analysis technique of optimization which is implemented using a Decision Tree Model (DTM) to optimize business investments. Another argument of this research is that in reality risk is better represented by two values: success and failure. This is achieved by finding the probability of an investment succeeding or failing. This gives investors a careful insight into what they stand to lose or gain if their investments fails or succeeds respectively.

2.1 Business Investment

Modern portfolio or investment selection theory usually deals with two opposing concepts – risk aversion and maximization of returns. The main point of modeling this problem is how the risk and asset profitability are defined and measured (Vercher, 2006). Portfolio selection deals with how to select and arrange a variety of assets in a way that best fulfills the investor's financial ambitions (Huang, 2008).

2.1.1 Business Investment Objectives

The investor's objectives are his/her investment goals expressed in terms of both risk and returns. The relationship between risks and returns requires that goals not be expressed only in terms of returns. Expressing goals only in terms of returns can lead to inappropriate practices such as the use of high-risk investment strategies or account "churning", which involves moving in and out of investments in an attempt to buy low and sell high (Reilly & Brown, 2002). There is an enormous amount of investment objectives made by investors and expressed in different ways, but they generally fall within three major categories:

1. Maximizing the total return on investment (ROI).
2. Limiting or minimizing risks associated with an investment. This is determined by the investor's risk tolerance levels.
3. To build an optimum portfolio through diversification.

2.1.2 Investment Principles

According to Investopedia, the online investment dictionary, investment principle is asset of guiding principles that inform and shape an individual's investment decision making process. Examples of investment principles or styles include:

1. Value Investing: This principle seeks relatively undervalued stocks believing they will eventually produce strong returns.
2. Fundamental Investing: Identifying companies with strong earnings prospects.
3. Growth Investing: This involves buying into companies that have promising emerging products or services that hold promising growth potential.
4. Socially-Responsible Investing: This investment principle looks for companies that adhere to certain set of moral and/or ethical business standards.
5. Technical Investing: This examines past market data to look for hallmark visual patterns in trading activity to make buy and sell decisions.
6. Contrarian Investing: This involves making investment decisions in direct opposition to the market majority (selling when others are buying); (Investopedia).

It should be noted that all the investment principles mentioned in this research study have their merits and demerits; and some investment principles offer more level of protection from huge amounts of investment losses or minimize investment risks than others.

2.1.3 Investment Variables

Various variables affect the quality of investment decision and optimization system. They include the following:

- **The Balance Sheet**: The balance sheet is a representation of the company's financial health. It is presented at a specific point in time, usually at the end of the fiscal (accounting) period, which could be a year, a quarter, or a month. It lists the assets that the company owns and liabilities the company owes to others (Fields, 2011). It is denoted as:

  Asset = Liabilities + Company's Equity
- **Expected Rates of Return on Investment:** Return on investment (ROI) is the benefit to the investor resulting from an investment of some resource. A high ROI means the investment gains compare favourably to investment cost (Wikipedia, Return on investment, 2015).
- **Risk and Risk Analysis:** Risk is an uncertainty that an investment will earn its expected rate of return (Reilly & Brown, 2002). When faced alternatives, investors have to carry out a comparative risk analysis to determine which investments holds the lowest risks.
- **Beginning Value of Investment:** This is the amount of money committed to an investment. For example, an investor bought stocks that cost NGN1 per share (Reilly & Brown, 2002).
- **End Value of Investment:** This refers to the amount of money an investor gets back at the end of an investment year. For example, an investor who spent NGN1 to purchase a unit of share gets a current share price of NGN1.50K at the end of an investment year (Reilly & Brown, 2002).
- **Measures of Historical Rates of Return:** This refers to rates of return for the period which an investor owns an investment (holding period). There are two categories of this measures, there are:
- **Holding Period Return (HPR):** The HPR computes the returns on an investment for a holding period and is denoted by

HPR = End Value of Investment / Beginning value of Investment.
HPR helps investors express the change in value of an investment. The following rules apply to the HPR:

- HPR value must always be zero or greater (i.e. it can never have negative value).
- A value greater than 1.0 reflects an increase in an investor's wealth.
- A value less than 1.0 reflects an investor suffered a decline in wealth

- **Holding Period Yield (HPY):** The HPY reflects the percentage return on an investment for a holding period. It is expressed in percentage and denoted by:
  HPY = (HPR – 1) * 100% (Reilly & Brown, 2002)
- **Annual Holding Period Return:** Annual HPR: The annual HPR is denoted by:

Annual HPR = $HPR^{1/n}$,
Where n = Number of years the investment is held (Reilly & Brown, 2002).

- **Mean Historical Returns:** This metric takes into consideration the mean rates of return for a single investment over a number of years (holding periods). This is a very important variable for the decision tree analysis.
  The decision tree requires and expected yield on an investment in order to compute the Expected Monetary Value (EMV) of an investment. The return on investment of an investment for a single year is not an accurate and justifiable parameter for such computation and the yield of several years cannot be supplied to the decision tree without causing complications. This implies that there is a need for a summary figure that indicates the investment's typical experience. Such a summarized figure can be derived by computing the mean annual rates of return on investment over some period of time. Reilly & Brown (2002) stated that, for a single investment, given a set of return (HPYs) the summary measures of return on investment is computed by the Arithmetic mean, the sum (∑) of annual HPYs and it is denoted as

$$AM = \sum HPY / n,$$

Where ∑HPY = the sum of the annual holding period yields
n = number of years the investment is held.

2.1.4 Data Quality and Requirements for Business Investment Optimization

Data quality is a multi-dimensional measurements of the adequacy of a particular datum or sets of data. In business, data quality is measured to determine whether or not data can be used as a basis for reliable business intelligence and for making organizational decisions. Data is a major resource required by all optimization models to make informed decisions that result in an optimized portfolio or investment. In order not to make misinformed and incorrect choices, one must ensure data's fitness to serve its purpose in the optimization context such as: *accuracy, completeness, relevance, consistency across data sources, and reliability.*

3. OPTIMIZATION

Optimization is one of the methods employed in the field of Operations Research as a tool to improve decision making and efficiency. An optimization model seeks to find values of the decision variables that optimize (maximize or minimize) an objective function among the set of all values for the decision variables that satisfy the given constraints (Winston, 2003).

3.1 Optimization Model

The Optimization Model consists of three sets of parameters as illustrated in Figure 3.1.

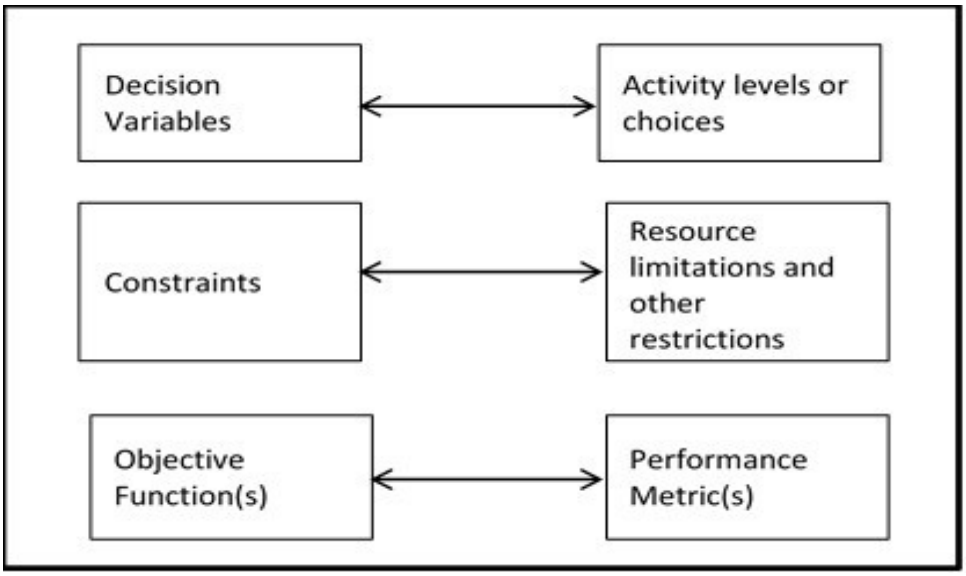


Fig 5.1: Box View of an optimization model

Optimization models consist of three components which are:

1. **Decision Variables:** These are elements which are under the control of the decision maker. They represent the activity levels or choices available to the decision maker.
2. **Constraints:** These are restrictions on the decision variables or resource limitations.
3. **Objective Function**: This refers to the performance metrics used to validate the performance of the outcome of the decision.

3.1.2 Decision Analysis

Decision Analysis is the discipline comprising the principle, theory, methodology and professional practice necessary to address important decisions in a formal manner. Decision Analysis includes many procedures, methods, and tools for identifying, clearly representing, and formally assessing important aspects of a decision. It also prescribes a recommended course of action by applying the maximum expected utility action axiom to a well-formed representation of the decision, and translates the formal representation of a decision and its corresponding recommendation into insight for the decision maker and other stakeholders (Wikipedia). Decision Analysis provides a framework and methodology for rational decision making when outcomes are uncertain (Hillier & Lieberman, 2010).

3.1.3 The Decision Analysis Process

The Decision Analysis process according to Ville (2006) is illustrated graphically in Figure 3.2

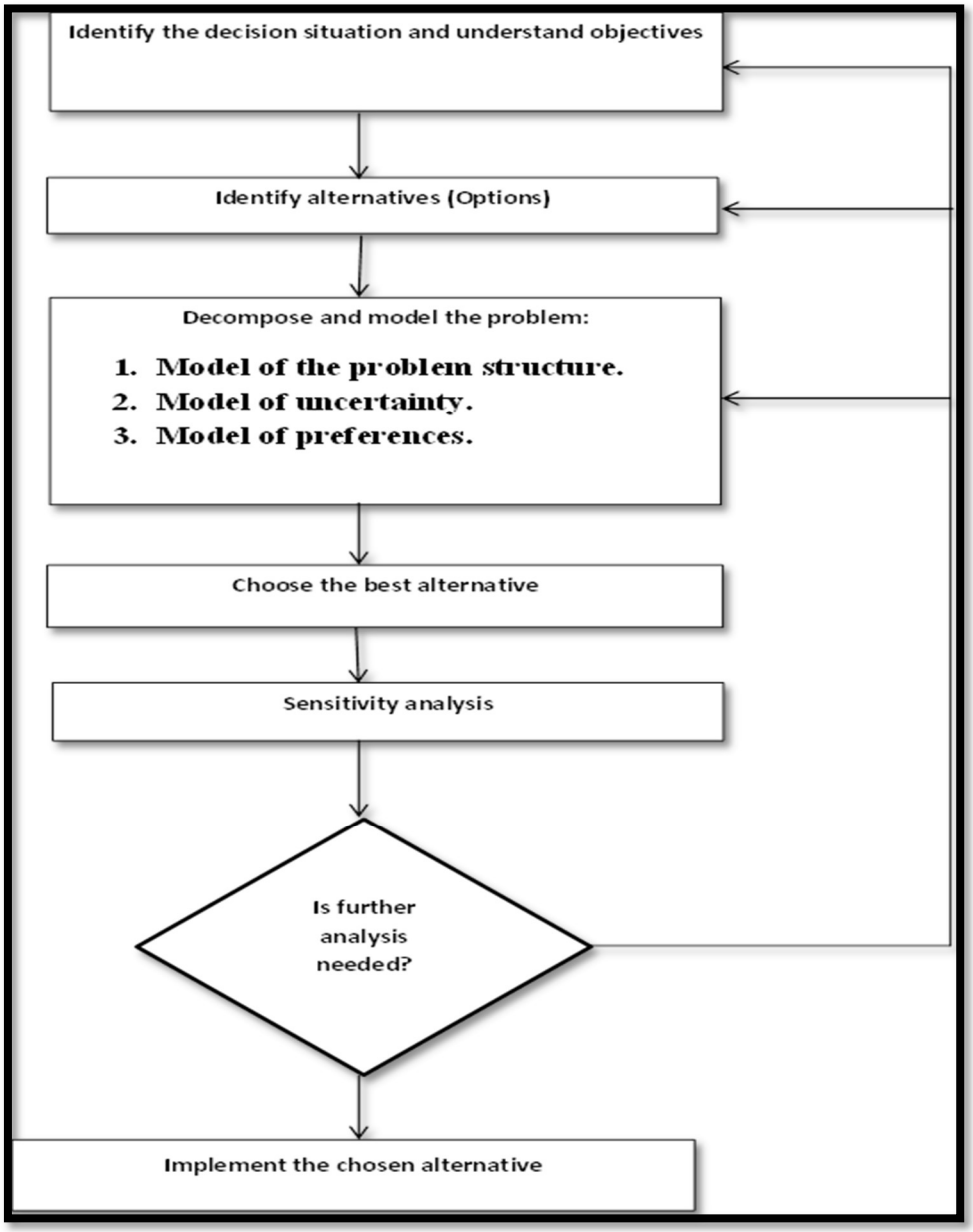


Fig 3.2: Decision Analysis process flowchart. Source: Ville (2006)

3.1.4 Elements of a Decision Problem

Clemen (2006) stated that in order to solve a decision problem it is important to identify the elements of the situation. These elements can be classified as follows: values and objectives, decisions to make, uncertain events, and consequences (outcomes).

- **Values and Objectives:** In this context values refer to the things that matter to the decision maker. An objective refers to a specific thing that the decision maker wants to achieve.
- **Decisions to Make (Alternatives):** This refers to the action the decision maker must make out of all the array of possible actions available to him or her. In some cases the optimal decision to make is to not take any action at all (i.e. wait till sometime in the future before taking an action.
- **Uncertain Events:** Decision problems can be complicated about what the future holds. Many important decisions have to be made without knowing exactly what will happen in the future or exactly what the ultimate outcome will be from a decision made today. The possible things that can happen in the resolution of an uncertain event are called outcomes.
- **Consequences:** After the last decision has been made and the last uncertain event has been resolved, the decision maker's fate is finally determined. It may be a matter of profit or loss as in the case of an investor. If a decision context requires consideration of multiple objectives, the consequence is what happens with respect to each of the objectives

3.1.5 Decision Rules

According to Ragsdale (2008), decision rules help a decision maker choose the best alternative and can be divided into two major categories namely: non-probabilistic methods and probabilistic methods.

- **Non-Probabilistic Methods:** Decision rules based on non-probabilistic methods assume that probability occurrence cannot be assigned to states of nature (i.e. uncertain events) in a decision problem. Non-probabilistic methods have two approaches to decision making which are:

1. ***The Maximax Decision Rule:*** This rule determines the maximum payoff for each alternative and then selects the alternative associated with the largest payoff. It should be noted that this approach does not guarantee that this payoff will occur and in some situations, the maximax decision leads to poor decisions
2. ***The Maximin Decision Rule:*** This is a more conservative approach to decision making which pessimistically assumes that nature will always be "against us" regardless of the decision we make. This decision rule can be used to hedge against the worst possible outcome of a decision.
   To apply the maximin decision rule, we first determine the minimum possible payoff for each alternative and then select the alternative with the largest minimum possible payoff (or the maximum of the minimum payoffs – hence the term "maximin").

- **Probabilistic Method:** Probabilistic decision rules can be used if the states of nature (uncertain events) in a decision problem can be assigned probabilities that represent the likelihood of occurrence. For decision problems that occur more than once, it is often possible to estimate these probabilities from historical data. However many decision problems represent one-time decisions for which historical data for estimating probabilities are unlikely to exist. I these cases, probabilities often are assigned subjectively based on interviews with one or more domain experts.

The probabilistic method has two approaches to decision making namely:

1. ***Expected Monetary Value Decision Rule:*** This selects the decision alternative with the largest expected monetary value (EMV). The EMV of alternative I in a decision problem is defined as
   $EMV_i = \sum r_{ij} P_j$

Where r_{ij} = the payoff for alternative I under the jth state of nature (uncertain event)
P_j= the probability of the jth state of nature (uncertain event)

2. ***Expected Opportunity Loss Decision Rule:*** This can use the probability of states of nature (uncertain events) to compute the expected opportunity loss (EOL), or the expected regret for each alternative in a decision problem. Therefore calculating the EOL is the same as that of the EMV, only that here we substitute regret values (opportunity loss for the payoffs. The EMV and EOL decision rules always result in the selection of the same decision alternative.

According to Ville (2006), for the application of probabilistic decision rules on a decision tree, the probabilities must satisfy either of the following requirements:

- *Probabilities must lie between 1 and 0:* Every probability (P) must be positive, and between 0 and 1 inclusive ($0 < P < 1$). This is a sensible requirement. In formal terms it simple means nothing can have more than a 100% chance of occurring or less than 0% chance.
- *Probabilities must add up:* Suppose outcomes are mutually exclusive (i.e. only one can happen, not both), the probability that one or the other occurs is then given the sum of the individual probabilities. Mathematically, we write $F(A_1 \text{ or } A_2) = P(A_1) + P(A_2)$ if A_1 and A_2 cannot both happen. For example, consider the stock market. Suppose there is a 30% chance that the market will go up and 45% chance that it will stay the same. It cannot do both at once, and so the probability that it will either go up or stay the same must be 75%.
- *Total probability must be equal to 1:* Suppose a set of outcomes is mutually exclusive and collectively exhaustive, this means that one (and only one) of the possible outcomes must occur. The probability for this set must sum to 1. Informally, if we have a set of outcomes such that one of them has to occur, then there is a 100% chance that one of them will indeed come to pass.

3.1.6 Decision Tree Model (DTM)

Wikipedia stated that "in computational complexity and communication complexity theories, the Decision Tree Model (DTM) is the model of computation or communication in which an algorithm or communication process is considered to be basically a decision tree. That is to say, a sequence of branching operations based on comparisons being assigned the unit computational cost. The branching operations are called "tests" or "queries". In this setting the algorithm in question may be viewed as a computational Boolean function $f: \{0, 1\}^n \rightarrow \{0, 1\}$, where the input is a series of queries and the output is the final decision. Every query is dependent on previous queries."

3.1.7 Decision Tree

A decision tree is a decision support tool that uses a tree-like graph or model of decisions and their possible consequences, including chance event outcomes, resource costs, and utility. Decision trees are commonly used in Operations Research, specifically in Decision Analysis to help identify a strategy most likely to reach a goal (Wikipedia). According to Ville (2006), decision trees are a simple, but powerful form of multiple variable analyses.

They provide unique capabilities to supplement, complement, and substitute for:

- Traditional statistical forms of analysis such as multiple linear regression.
- A variety of data mining tools and techniques (such as neural networks).
- Recently developed multidimensional forms of reporting and analysis found in the field of business intelligence.

3.1.8 Some Rule of Thumb for Decision Tree Model (DTM)

Ville (2006) believed that for a decision tree to be properly modelled that the following rules should be adhered:

1. Understand the differences between categorical and continuous data. Categorical data such as zip codes have a numeric form with many values that can look like continuous data, but are actually categories. Consider clustering categories together outside of the decision tree. It might be possible to cluster categories together with respect to a target variable.
2. Categorical targets with more than two values are extremely difficult to interpret. Rework multi-category targets into a 1-of-Node scheme. With 1-of-N coding, each distinct category becomes a binary yes-no/on-off outcome in a new input. There are as many binary inputs as categories in the original multi-category input.
3. Dates can be continuous fields, but might need to be changed to Julian format. It is useful to compute time intervals, such as length of time as a customer, before beginning analysis.
4. There can be other time and distance measures; these need to be calculated and verified before analysis.
5. Try to avoid information loss; higher levels of measurement contain more information than lower levels, so actual income is preferable to income wages.
6. If you are working with variables that are expressed along a scale, for example 1, 2, 3, and so on, then you might find it easier to express all scales in the same direction.
7. Multiple response items might need to be treated with care. For example, if you have a list of products that are purchased, then each product might need to be totalled separately, and a total number of products purchased might need to be calculated. In this, multiple response items within each unit of observation might need to be summed to create an analysis data set.
8. Do not confuse missing information with a missing value because this is not always the same. For example, quantity purchased can be blank for a given day or a given product type if the customer did not purchase that product.
9. It might be necessary to pivot records, particularly if you want to compute purchase quantities for given products. The product purchase tends to be one-line-per-purchase records with purchase details and a customer number. The purchase details need to be summarized through a pivoting operation, such as PROC TRANSFORM. Then the aggregates are attached to the record (typically using customer number as a key).

3.1.9 Decision Tree Growing Sub-Processes:

Ville (2006) stated that the decision tree growing process can be broken down into a number of sub-processes as shown in Figure 3.3

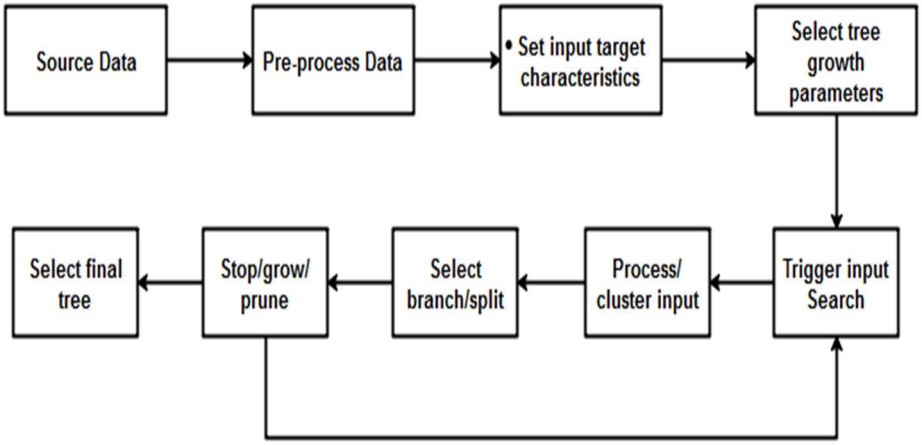


Fig 3.3: Illustration of sub-processes in growing a decision tree. Source: Ville (2006)

4. RESEARCH METHODOLOGY

This refers to the principle(s) used to carry out this research work. It analyses the principle(s) of the model employed by this research work. The research methodology employed in this research is the Decision Analysis method/technique. Decision Analysis method is a systematic, quantitative and visual approach to addressing and evaluating important investment alternatives (options). This method evaluates complex alternatives (options) in terms of values (in this case monetary values) and uncertainty. It provides an understanding of how the alternatives differ from each other and the weighted values for each outcome in order to determine the outcome that meets the investment optimization objective. The Decision Analysis technique will be implemented with a Decision Tree Model (DTM).

4.1.1 Analysis

Analysis is the process of investigating any current system(s), identifying improvement opportunities, and developing a concept for the new system. The objective of analysis is to understand and document in detail the needs and the processing requirements of the new (proposed) system.

4.1.2 Analysis of Existing System (Framework)

Mathematical Optimization techniques has been the most used technique in investment or portfolio optimization. A very popular implementation of the Mathematical Optimization technique is the Nonlinear Programming (NLP) model. According to Hillier & Lieberman (2010), it now is common practice for professional managers of large stock portfolios to use computer models based partially on nonlinear programming to guide them. Because investors are concerned about both the *expected return* (gain) and the *risk* (possibility of loss) associated with their investments, nonlinear programming is used to determine a portfolio that, under certain assumptions, provides an optimal trade-off between these two factors. This approach is based largely on path-breaking research done by Harry Markowitz and William Sharpe that helped them win the 1990 Nobel Prize in Economics.

Suppose that an investor is considering including *n* stocks (securities) to a portfolio, the Nonlinear Programming (NLP) model can be formulated for this problem as follows:

$$R(x) = \sum_{j=1}^{n} u_j x_i \ > L$$

$$\sum_{j=1}^{n} P_j X_i \ < B$$

And $x_i \ \geq 0, for\ j = 1, 2, \ldots \ldots . . n$

Where:
V(*x*) = Variance of the total return from the entire portfolio
R(*x*) = Expected Value. This could be used as a constraint in the NLP model
i = Represents the alternating stocks of interest when $i \neq j$
j = Represents stocks of interest (*j* = 1, 2, …..*n*).
σ_{ij}= The covariance of the return on one share of each stock *i* and stock *j*
σ_{jj}= Variance of the on each share of stock *j*. This measures the risk of this stock
σ_j= Standard Deviation of each share of stock *j*
μ_j = Mean of the return on each share of stock *j*
x_i=Number of shares of stock *i* when $i \neq j$
x_j= Decision variables representing number of shares of stock to be included in the portfolio. (*j* = 1, 2, …..*n*).
P_j= Price for each share of stock *j*.
L = Minimum acceptable expected return
B = Amount of money budgeted for the portfolio

One drawback of this formulation is that it is relatively difficult to choose an appropriate value for *L* for obtaining the best trade-off between *R*(**x**) and *V*(**x**). Therefore, rather than stopping with one choice of *L*, it is common to use a *parametric* (nonlinear) programming approach to generate the optimal solution as a function of *L* over a wide range of values of *L*. The next step is to examine the values of *R*(**x**) and *V*(**x**) for these solutions that are optimal for some value of *L* and then to choose the solution that seems to give the best trade-off between these two quantities. For example, given the data of stocks in Table 4.1

Table 6.1: Data of stocks

Stock	Expected Return (%)	Risk (%) (Standard Deviation)	Pair of Stocks	Joint Risk per Stock (Covariance)
1	21	25	1 and 2	0.040
2	30	45	1 and 3	–0.005
3	8	5	2 and 3	–0.010

What combination of stocks in a portfolio will minimize the risk subject to achieving at least an 18% expected return? The formula for this problem is formulated as follows:

$$\text{Minimize Risk} = (0.25Z_1)^2+(0.45Z_2)^2+(0.05Z_3)^2+2(0.04)Z_1Z_2+2(-.005)Z_1Z_3 +2(-0.01)S_2S_3$$

subject to: $(21\%)Z_1 + (30\%)Z_2 + (8\%)Z_3 \geq 18\%$

$Z_1 + Z_2 + Z_3 = 100\%$

and

$Z_1 \geq 0,\ Z_2 \geq 0,\ Z_3 \geq 0.$

The result from the computation of the nonlinear programming model is shown in Table 6.2.

Table 6.2: Result of Nonlinear Programming Model Analysis

	Stock 1	**Stock 2**	**Stock 3**	
Expected Return	21%	30%	8%	
Risk (Stand. Dev.)	25%	45%	5%	
Joint Risk (Covariance)	**Stock 1**	**Stock 2**	**Stock 3**	
Stock 1		0.040	-0.005	
Stock 2			-0.010	
Stock 3				
	Stock 1	**Stock 2**	**Stock 3**	**Total**
Portfolio	40.20%	21.70%	38.10%	100%
	Portfolio			**Minimum Expected Return**
Expected Return	18.00%		>=	18.00%
Risk (Variance)	0.0238			
Risk (Stand. Dev.)	15.40%			

5. SYSTEM DESIGN

System design is the process of deciding how the system will operate. This phase of the system development process uses information collected in the analysis phase to accomplish the logical design of the system. The objective of system design is to design the solution system based on the requirements defined and decisions made during analysis.

5.1 Proposed System Design

One of the objectives of this research work is to implement a business investment optimization system using a decision tree model also known as BIOS-DTM. This section examines the design of the BIOS-DTM decision tree.

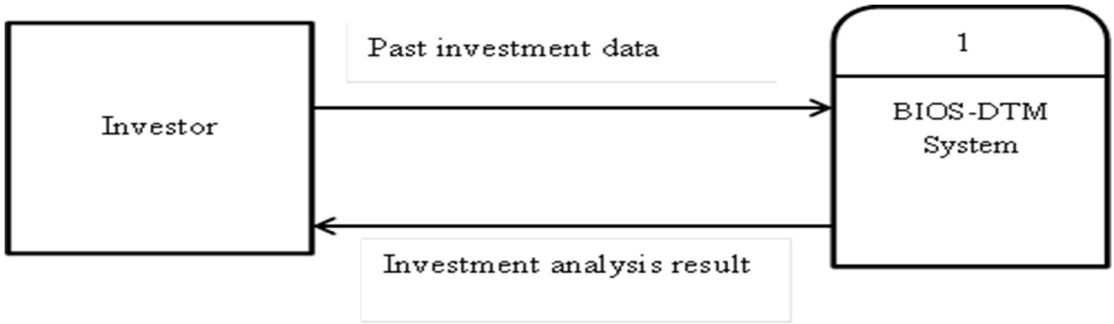


Fig 5.1: Data Flow Diagram (Context Diagram) of the BIOS-DTM system

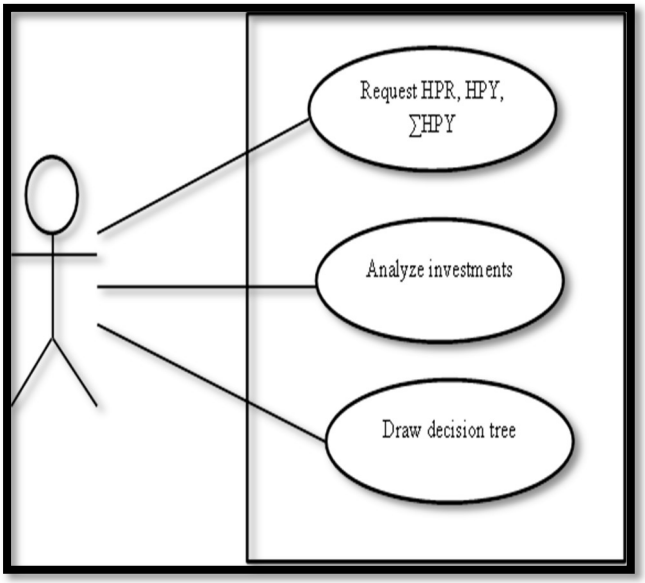


Fig 5.2: Use Case Diagram of the BIOS-DTM system

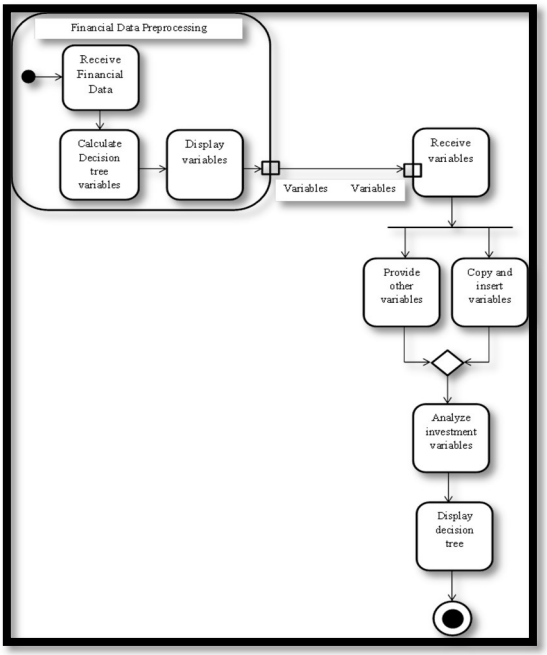


Fig 5.3: Activity Diagram for BIOS-DTM system

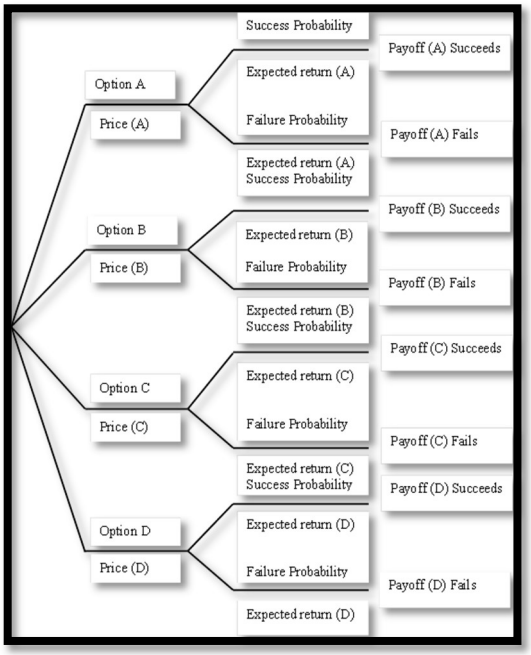


Fig 5.4: Decision tree User Interface Design

5.2 Architecture of the Proposed System

This is a description of how software systems are organized. It refers to the high level structures of software systems, and the documentation of these structures.

The architecture of the BIOS-DTM system is shown in figure 5.5.

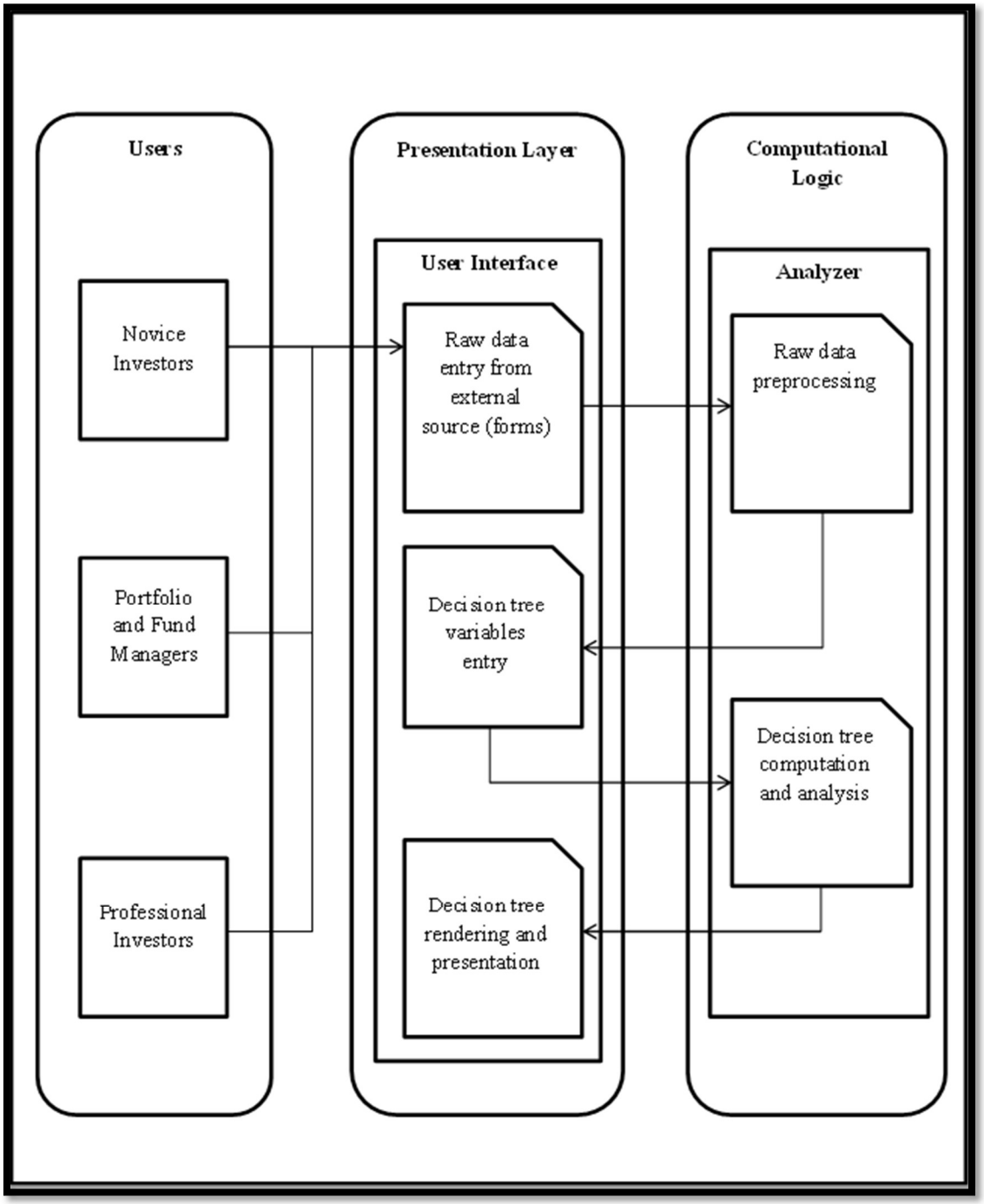


Fig 5.5: BIOS-DTM architecture

5.3 Description of Key Components

- Option(s): This represents the investment alternatives that an investor has to choose from.
- Price: The cost of purchasing an investment, in this case, shares.
- Probability: The likelihood that an investment will succeed or fail.
- Cash Flow: This is the amount of money an investment is expected to accrue over a certain period of time after it has been purchased.
- Payoff(s): The amount of money gained or lost after paying for an investment. It is usually the subtraction of the cost of an investment from the expected return on the investment.
- Expected Monetary Value (EMV):The total of the weighted outcomes (payoffs) associated with a decision, the weights reflecting the probabilities of the alternative events that produce the possible payoff.

5.3.1 Data Source

The main source of data for this research work was secondary data. Data was gathered from secondary sources such as financial annual reports, financial websites, and shareholders' annual general meeting brochure. A form was constructed and used to extract specific sets of data from the various secondary data sources.

6. IMPLEMENTATION PROCEDURE

The system design of this research work was implemented in Java. It was chosen as the implementation language for the following reasons:

- Familiarity with the language.
- Its support for cross-platform development.
- Its support for dynamic memory use
- Java is an easy-to-main program.
- Java programs are secure.

The tools used for the development of the BIOS-DTM system include Microsoft Word 2010 and Netbeans 8.0. The system designs in this research were created with Microsoft Word 2010 and Netbeans 8.0 was used to implement the Java program for this research. Netbeans 8.0 was chosen for this research work for the following reasons:

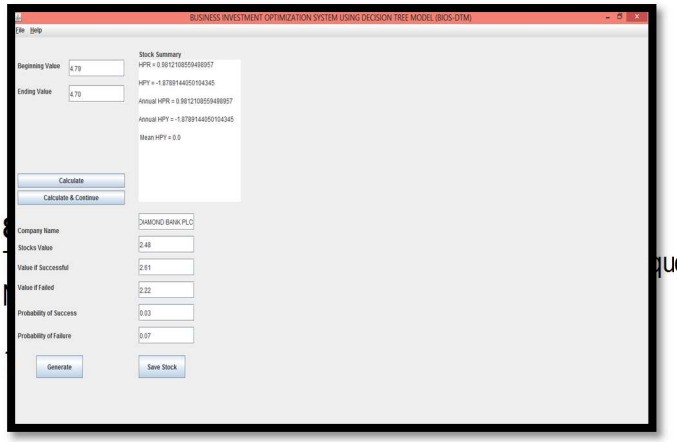
- Fast and smart code editing.
- Easy and efficient program management.
- Rapid User Interface development (GUI design tool).
- The presence of debugger and profiler features for static code analysis.
- Its cross platform support.

The program in this research was implemented on the Windows operating system platform because it is the most widely used operating system in Nigeria.

6.1 Algorithm

Start Program
Enter Beginning Value and End Value
If number of years is applicable
    Enter number of years
    Compute HPR, HPY, HPRa, HPYa
    Print HPR, HPY, HPRa, HPYa on screen
Else compute HPR, HPY, HPRa, HPYa, and meanHPY
Repeat steps 2 and 4 until data is exhausted
Print HPR, HPY, HPRa, HPYa and meanHPY on screen
Read meanHPY
Repeat steps 2 to 7 for periods of investment decline in value
Enter a Company Name
Enter stock price
Enter meanHPY for success
Enter meanHPY for failure
Enter probability for investment succeeding
Enter probability for investment failing
Compute payoffs
Compute EMVs
Determine the investment that satisfies the objective function
Save the results in an Array List
Repeat the whole process 4 times from step 2 to
Generate and print Decision Tree
Print the EMV of the investment that satisfies the objective function
End program

6.2 Sample Implementation Input Snapshot



6.1 Sample Implementation Output Snapshot

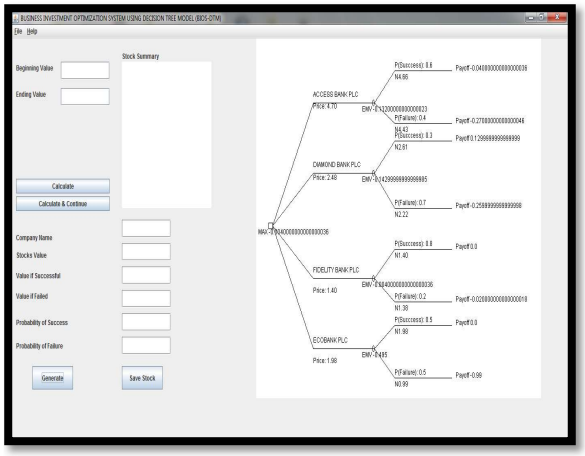


Fig 6.2: Financial services Sector Result display screen of the BIOS-DTM

Table 6.1: Four companies each are selected from five sectors from the economy for study

S/N	SECTOR	COMPANY	PRICE	CF(S)	CF(F)	P(S)	P(F)	PO(S)	PO(F)	EMV	O.S
1	FINANCIAL SERVICES	ACCESS BANK PLC	4.70	4.66	4.43	0.6	0.4	-0.04	-0.27	-0.132	
2		DIAMOND BANK PLC	2.48	2.61	2.22	0.3	0.7	0.13	-0.26	-0.143	
3		FIDELITY BANK PLC	1.40	1.40	1.38	0.8	0.2	0.00	-0.02	-0.004	*
4		ECOBANK PLC	1.98	1.98	0.99	0.5	0.5	0.00	-0.99	-0.495	
5	AGRICULTURE	ELLAH LAKES PLC	4.26	4.26	2.13	0.5	0.5	0.00	-2.13	-1.065	
6		FTN COCOA PROCESSORS PLC	0.50	0.50	0.25	0.5	0.5	0.00	-0.25	-0.125	
7		LIVESTOCK FEEDS PLC	1.52	1.44	1.29	0.8	0.2	-0.08	-0.23	-0.11	
8		OKUMO OIL PALM PLC	26.33	28.38	23.84	0.8	0.2	2.05	-2.49	1.142	*
9	OIL AND GAS	CONOIL PLC	27.41	27.41	13.71	0.5	0.5	0.00	-13.7	-6.85	
10		MOBIL OIL NIGERIA PLC	138.99	138.99	69.50	0.5	0.5	0.00	-69.49	-34.745	
11		OANDO PLC	6.19	7.71	2.90	0.2	0.8	1.52	-3.29	-2.328	*
12		TOTAL NIGERIA PLC	143.00	143.00	71.50	0.5	0.5	0.00	-71.5	-35.75	
13	CONSUMER GOODS	7UP BOTTLING COMPANY PLC	184.00	183.43	179.51	0.5	0.5	-0.57	-0.49	-2.53	
14		CADBURY NIGERIA	20.00	19.46	19.00	0.8	0.2	-0.54	-1.0	-0.632	
15		CHAMPION BREWERY PLC	3.80	4.01	3.43	0.8	0.2	0.21	-0.37	0.094	*
16		DANGOTE SUGAR REFINERY PLC	6.20	6.34	5.89	0.6	0.4	0.14	-0.31	-0.04	
17	HEALTHCARE	AFRIK PHARMACEUTICALS PLC	0.50	0.50	0.25	0.5	0.5	0.00	-0.25	-0.125	
18		EKOCORP PLC	3.72	3.72	1.86	0.5	0.5	0.00	-1.86	-0.93	
19		EVANS MEDICAL PLC	0.50	0.50	0.25	0.5	0.5	0.00	-0.25	-0.125	
20		FIDSON HEALTHCARE PLC	3.00	3.01	2.83	0.8	0.2	0.001	-0.17	-0.026	*

From Table 6.1 it can be seen that four companies each are selected from five sectors from the economy for study. The price column lists the end value of stocks which will be the beginning value for the next day trade. **Cash flow for success, CF(S)** is the appreciation value expected from an investment. In this work it has been derived by computing the average for a set of stock end values that indicate appreciation found in the collection of stock values (both gains and losses) over a period of time (at the discretion of the investor) for the analysis. Cash flow for failure, CF(F) is the depreciation value expected from an investment.

The rationale for determining the values of this column is based on finding the worst-case scenario of the loss an investor could suffer from an investment. It was derived by picking the highest depreciation of an investment from a range of depreciation values found in a collection of stock values (both gains and losses) over a period of time (at the discretion of the investor) and subtracting the net loss suffered by an investment from it.

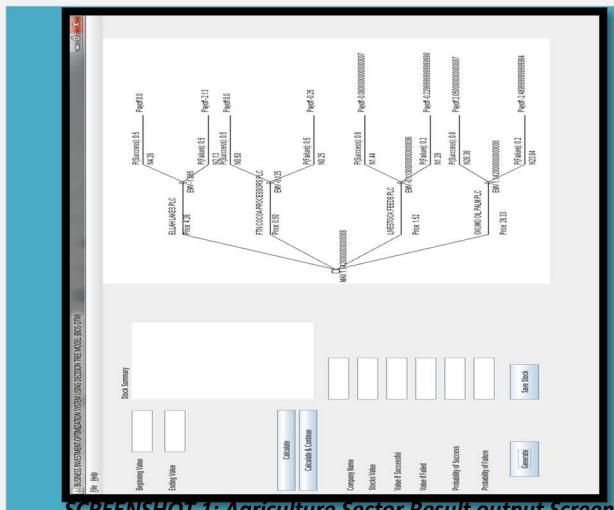
Probability of an investment succeeding, P(S) is derived by picking the number of gains found in a collection of stock values (both gains and losses) over a period of time.

Probability of an investment failing, P(F) is derived by picking the number of losses found in a collection of stock values (both gains and losses) over a period of time.

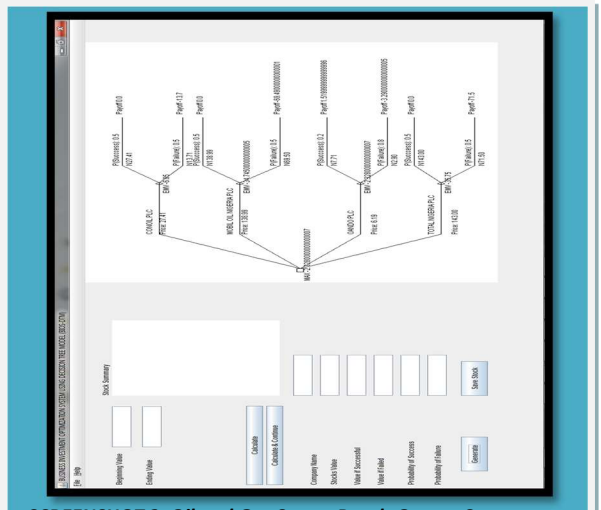
Payoff for investment success, PO(S) is derived from subtracting the investment price (i.e. cost) from the expected cash flow for success.
Payoff for investment failure, PO(F) is derived from subtracting the investment price (i.e. cost) from the expected cash flow of failure.

Expected monetary value (EMV) is used to determine to the objective function (maximize profit) of an investment. It is derived by summing the product of the probability for success and cash flow for success, and the product of the probability of failure and cash flow of failure.

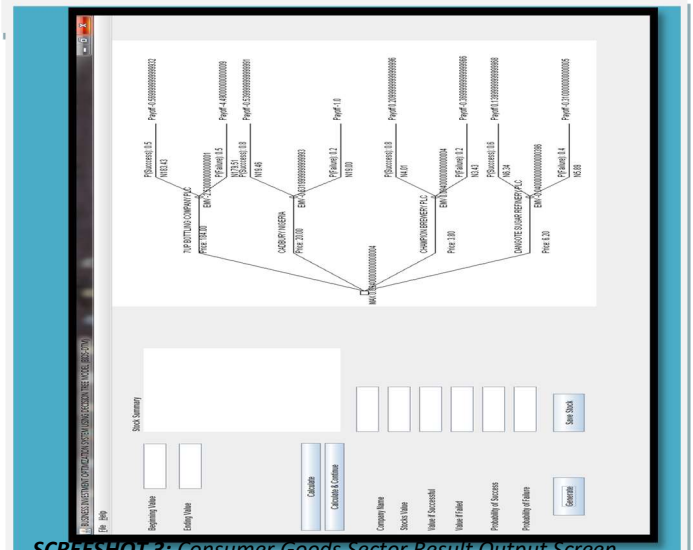
The optimal solution column is used to indicate the investment that actually satisfies the objective function, in this case maximize profit.



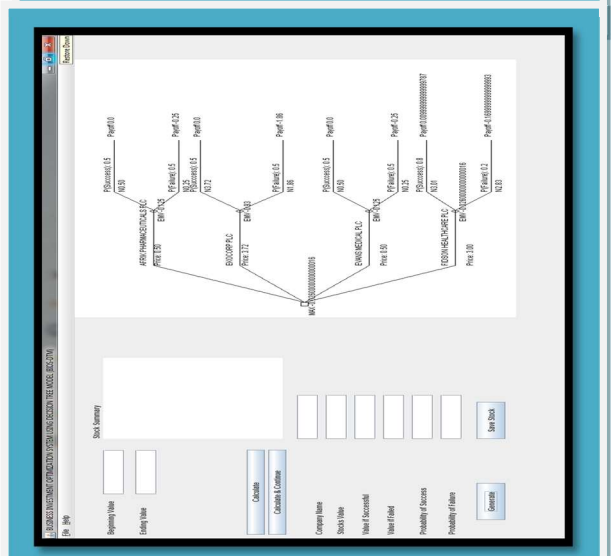
***SCREENSHOT 1:** Agriculture Sector Result output Screen*



***SCREENSHOT 2:** Oil and Gas Sector Result Output Screen*



***SCREESHOT 3:** Consumer Goods Sector Result Output Screen*



***SCREENSHOT 4:** Healthcare Sector Result Output screen*

7. DISCUSSION OF RESULTS

Examining the decision tree, it is observed that there is no correlation between the payoff of success and the EMV. This is to say an investment with the highest payoff for success does not always mean it satisfies the objective function to maximize profit.

There is no correlation between the size of the cash flow for success and the EMV. The investment with the highest cash flow for success does not always have the highest EMV which satisfies the objective function to maximize profit. In the Decision Tree Model (DTM) the investment that satisfies the objective function to maximize profits (i.e. EMV with the highest value) does not always translate to the investment with the lowest risk. It some cases it could be the investment with the highest risk of greater loss.

8. CONCLUSIONS

This entire research work has been focused on the Decision Tree Model (DTM), its underlying concepts, and its implementation as a model for optimizing business investments. Attempts have been made to investigate and analyse business investments, design, implement, test, and evaluate a business investment optimization system using a Decision Tree Model (DTM). The decision tree implemented in this research work has been used to demonstrate the optimization of business investments. It can be seen that the decision tree can pictorially show investment analysis for every possible investment decision. In this research work a Decision Tree Model (DTM) for the optimization of business investment has been presented. The Decision Tree Model (DTM) (DTM) developed is only applicable to portfolios containing only stocks (shares), provided data for return and risk forecast are available.

This research has shown that only the Decision Tree Model (DTM) takes into consideration the potential loss associated with an investment if an investment actually fails to meet the investor's objective of maximizing profit.
Some of the insights gained from this research were:

- The Decision Tree Model (DTM) is the easiest optimization model to understand in business investment optimization.
- It is one of the very few optimization models that consider the transaction cost of an investment.
- It can either be used on its own or combined with other Mathematical Optimization models for business investment optimization.
- It is a very efficient model for multiple variable analyses.

9. RECOMMENDATIONS

This research study recommends the following to investors or potential investors:

- Investors should adopt the value investing principle in order to avoid buying over-valued stocks.
- Investors should calculate the capital intensity of their potential investment since the value investing principle fails to put this factor into consideration. This is because a low capital intensive business is more profitable than a high capital intensive business even if/when they earn the same gross profit.
- Investors that are extremely averse to risk should apply the more conservative "maximin" decision rule to the decision tree analysis.

10. SUGGESTION FOR FURTHER RESEARCH

While this research provides the basic framework for simplifying business investment optimization, more work is needed in several areas. These areas include the following:

- The decision tree optimization model can be refined to provide probabilistic and non-probabilistic modes of decision making.
- The Decision Tree Model (DTM) can be enhanced and extended to optimize other business investment vehicles such as bonds, real estate, treasury bills, etc.
- The Decision Tree Model (DTM) can be extended to approach business investment optimization from the perspective of several investment principles, one at a time.

REFERENCES

Batini, C., & Scannapieca, M. (2006). *Data Quality: Concepts, Methodologies and Techniques.* Springer.

BUSETTI, F. R. (2000, June). Metaheuristic Approaches to Realistic Portfolio Optimization. South Africa: University of South Africa.

Byrd, D. M., & Cothern, R. C. (2000). *INTRODUCTION TO RISK ANALYSIS: A Systematic Approach to Science-based Decision Making.* Government Institutes.

Chang, T. J., Meade, N., Beasley, J. E., & Sharaiha, Y. M. (1998). Heuristics for cardinality constrained portfolio optimization. The Management School, Imperial College, London.

Clemen, R. T. (1996). *Making Hard Decisions: An Introduction to Decision Analysis* (2nd ed.). Duxbury Press.

Fang, Y., Lai, K. K., & Wang, S. (2008). Fuzzy Portfolio Optimization: Theory and Methods. Springer.

Fields, E. (2011). *The Essentials of Finance and Accounting for Non-Financial Managers* (2nd ed.). Amacom.

Ghate, A. (2009, July 31). Introduction to Optimization Models. Seattle, USA: University of Washington.

Gilli, M. (2004, September). An Introduction to Optimization Heuristics. Department of Public and Business Administration, University of Cyprus.

Goodwin, P., & Wright, G. (2004). *Decision Analysis for Management Judgment* (3rd ed.). John Wiley & Sons, Ltd.

Hillier, F. S., & Lieberman, G. J. (2010). *Introduction to Operations Research* (9th ed.). New York: McGraw-Hill.

Investopedia. (n.d.). *Investment Philosophy*. Retrieved October 01, 2015, from Investopedia: http://www.investopedia.com/terms/investment-philosophy.asp

Konno, H., & Yamakazi, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo Stock Market. *Management Science 37*.

Koochakzadeh, N. (2013, March). A Heuristic Stock Portfolio Optimization Approach Based on Data Mining Techniques. Department of Computer Science Calgary, Alberta.

Mansourfar, G. (2013). Econometrics and Metaheuristic Optimization Approaches to International Portfolio Diversification. *Iranian Journal of Management Studies (IJMS)*, 45-75.

Merriam-Webster. (n.d.). *Investment*. Retrieved October 01, 2015, from Merriam-Webster: http://i.word.com/idictionary/Investment

Papahristodoulou, C., & Dotzauer, E. (2004). Optimal Portfolios Using Linear Programming Models. *Journal of the operational Research*, 55.

Ragsdale, C. T. (2008). *Spreadsheet Modelling & Decision Analysis: A Practical Introduction to Management Science* (5th ed.). Thomson South-Western.

Reilly, F. K., & Brown, K. C. (2002). *Investment Analysis and Portfolio Management* (7th ed.). South-Western College Publlishers.

Sebastian-Coleman, L. (2013). *Measuring Data Quality for Ongoing Improvement: A Data Quality Assessment Framework.* Morgan Kaufmann.

Vercher, E. (2006). Optimization in Finance: Portfolio selection models. *Iberian Conference in Optimization*, (p. 6). Coimbra.

Ville, B. d. (2006). *Decision Trees for Business Intelligence and Data Mining Using SAS Data Miner.* Cary, NC, USA: SAS Publishing.