Marking Scheme for Question 2

(a) Explain in details how Genetic Programming works. [30%]

This is book work. 5% will be awarded to each of the following points:

- Genetic Programming is basically a *search algorithm*.
- To apply genetic programming, one needs a *tree representation* of candidate solutions.
- To apply genetic programming, one needs a way to *evaluate* candidate solutions.
- In genetic programming, one maintains a *population* of candidate solutions.
- In genetic programming, a *reproduction* operator must be defined, which picks candidate solutions according to their fitness
- In genetic programming, the *crossover* operator involves exchange of sub-branches
- In genetic programming, the *mutation* operator involves modification of sub-branches
- *Exploitation* is an important element in genetic programming. This may be
- *Exploration*, which involves *randomness*, is an important element in genetic programming
- An appropriate *balance* between exploration and exploitation is important in the success of genetic programming and any other evolutionary algorithms.

(b) Explain, with the help of an example, how a forecasting strategy can be represented for Genetic Programming to use. [30%]

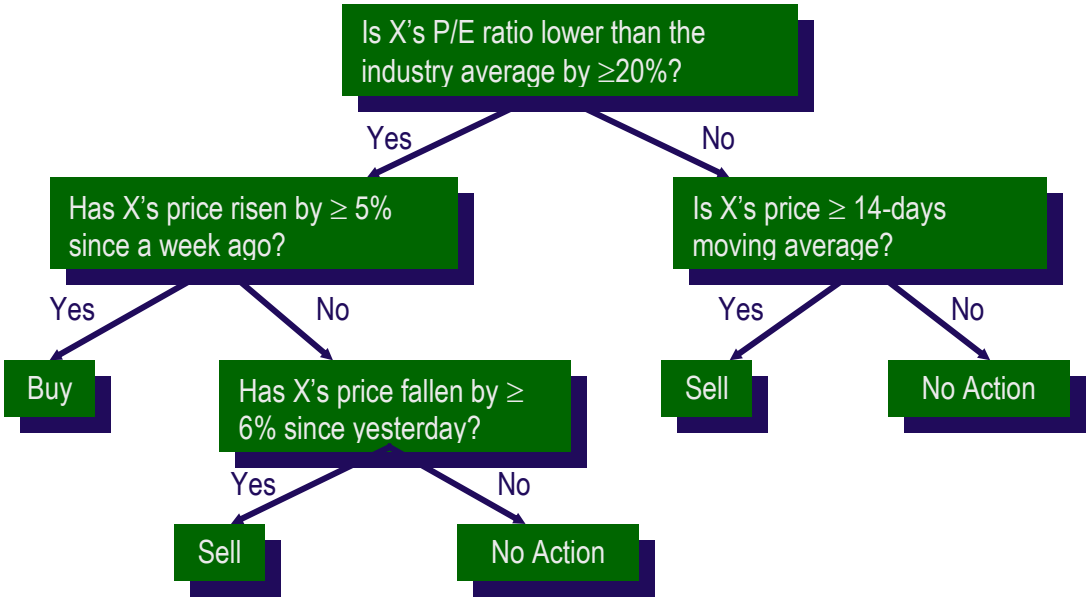
We need to represent a forecasting strategy in trees form. A tree can be represented in BNF form. A grammar in BNF form consists of functions and terminals.
[Students are not required to name the terminology "BNF form", but they need to be able to describe exactly how a grammar looks like.]

Here is an example of a grammar:

<Tree> ::= "If-then-else" <Condition> <Tree> <Tree> | Decision
<Condition> ::= <Condition> "And" <Condition> | <Condition> "Or" <Condition> | "Not" <Condition> | Variable <RelationOperation> Threshold
<RelationOperation> ::= ">" | "<" | "="

Terminals:
Variable is an indicator / feature
Decision is an integer, "Positive" or "Negative" implemented
Threshold is a real number

Following is an example tree (details of the terminals skipped here):



(c) Explain how the strategy that you have proposed above can be evaluated. [10%]

Strategies can be evaluated using past data. This will give each strategy a score, based on how accurate they forecast.

(d) Would the strategies trained with data from a bull market using genetic programming be useful in a bear market? Justify your answer carefully. [10%]

Students may answer yes or no, as long as they provide the right justification.
Yes, if one manages to find patterns in a bull market that capture price movements in the market.
No, if the nature of the market has changed fundamentally, and therefore pattern found in the bull market never appear in the bear market.

(e) Would evolutionary computation have been used by a University student on financial forecasting 30 years ago? Please justify your answer carefully. [10%]

The answer is probably not. 5% for each of the following points:

- There was not as much data.
- Computers were not fast enough to run evolutionary computation in large scale.
- Knowledge in evolutionary computation was limited.

(f) What will happen to the prices of an equity if everyone in the market uses the same algorithm to determine whether to buy or sell? [10%]

The forecast will be self-fulfilling: if everybody decides to buy (sell), then the price will indeed go up (down).