Software Measurement

In the field of software engineering, software measuring is critical. People working in the IT industry are continually confronted with new technology as well as other factors such as competitive marketplaces. They must also consider the product's dependability, stability, and testing while doing so. As a result, software measurement is required to help with these concerns. Although software measurement cannot address the challenges that IT professionals confront on its own, it may aid in clarifying and focusing their knowledge of issues. Sequential measurements of quality characteristics of goods and processes may also offer an efficient basis for launching and monitoring process improvement efforts provided software measurement is done effectively. Measurement of software engineering processes may be done in a variety of methods. Lines of code, balance scorecard, and instruction path length will all be discussed in my article.

Lines of Code (LOC)

This is sometimes referred to as source code lines (SLOC). This is a software statistic that counts the amount of text lines in the source code of a computer program to determine its size. SLOC is mostly used by programmers to determine how much work remains until the software is completed. This statistic is also used to assess the program's productivity and maintainability once it has been developed.

There are two primary types of SLOC measures:

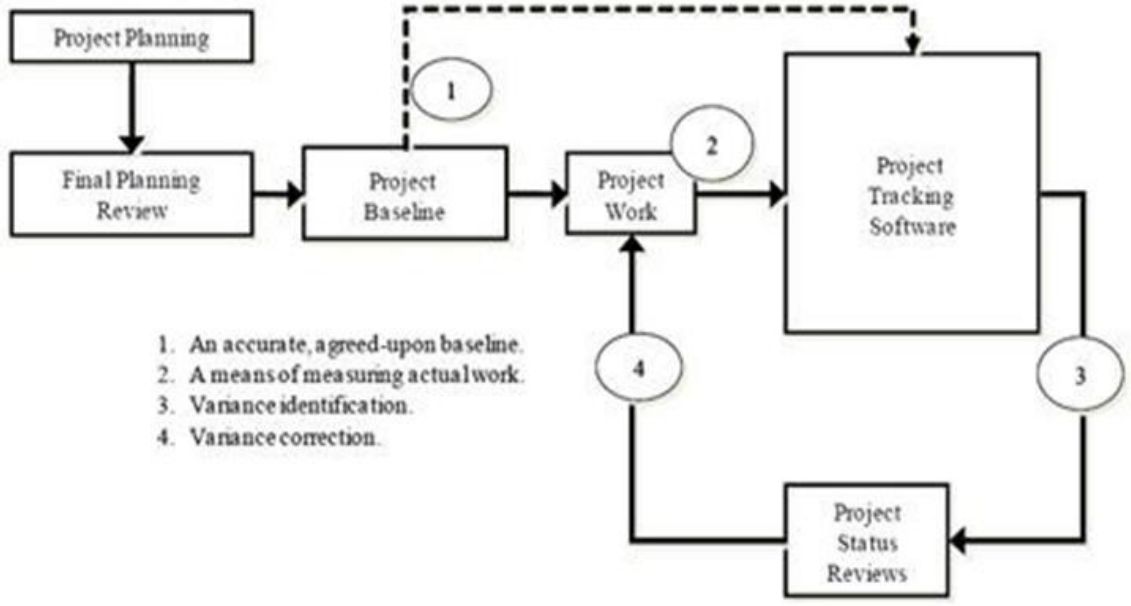
a) Physical SLOC (LOC) - It is a count of lines in the text of the program's source code, including comment lines and blank lines, unless a section's lines of code include more than 25% blank lines.

b) Logical SLOC(LLOC) - The number of "statements" is measured, although the definitions vary depending on the computer language. For example, the number of statement-terminating semicolons is a basic logical LLOC metric for C-like programming languages.

Making tools to assess physical SLOC is a lot simpler. Physical SLOC definitions are also simpler to describe than logical SLOC definitions.

Balance Scorecard

A Balance Scorecard is an example of a cybernetic control used in plan implementation management. A self-correcting feedback loop is used in cybernetic control.



Instruction Path Length

A program's instruction route length is the number of machine code instructions required to complete a segment. The total instruction path length is a measure of the algorithm's performance on a certain piece of PC hardware. A basic standard conditional instruction's route length is usually believed to be equal to 2.

The comparison operation is performed by one instruction, and the branching operation is performed by the second instruction whether the condition is true or false. The amount of time it takes to complete each command is often employed as a measure of relative performance rather than absolute and definite performance.

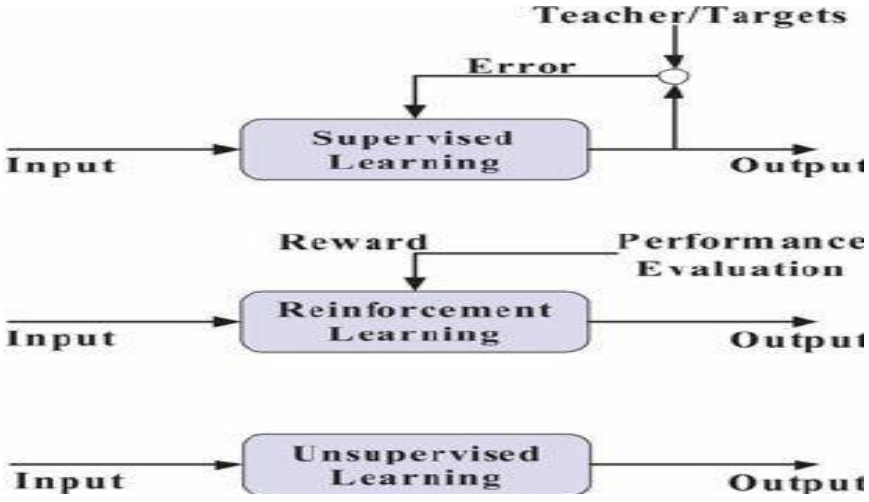
Machine Learning

What is Machine Learning, and how does it work? "Machine learning is an artificial intelligence (AI) application that allows systems to learn and improve from experience without having to be explicitly programmed."

Machine learning focuses on teaching a computer program to develop and learn on its own rather than relying on pre-programmed instructions to tell it what to do.

Machine Learning is divided into 3 types of algorithms:

1) Unsupervised Learning
2) Supervised Learning
3) Reinforcement Learning

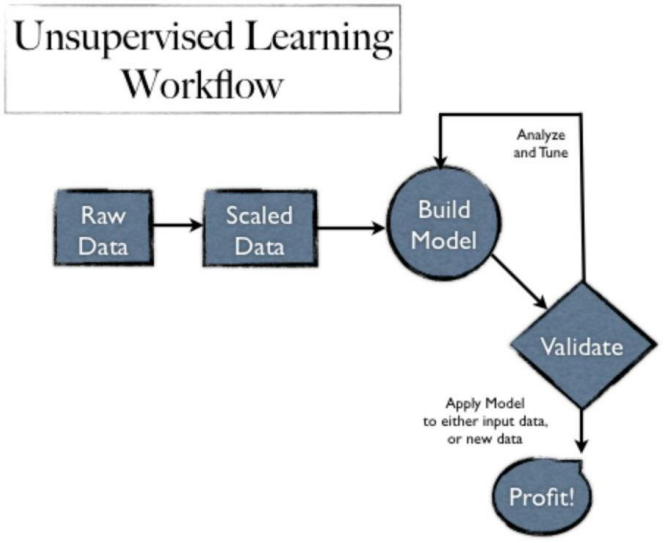


Unsupervised Learning

Unsupervised learning is a sort of machine learning method that is used to make conclusions from datasets containing unlabeled input data.

The learning algorithm creates an inferred function to generate predictions about the output values based on the examination of a known training dataset. After enough training, the system can offer objectives for any new input. The learning algorithm may also compare its output to the proper, intended output and detect faults, allowing the model to be modified as needed. 10

The most common sort of unsupervised learning is cluster analysis. It is used to discover hidden patterns and groupings in data. The clusters are represented using a similarity measure based on metrics like Euclidean or probabilistic distance.

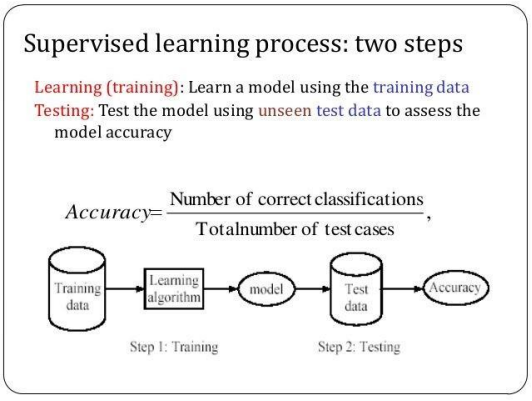


Supervised Learning

In practical machine learning, supervised learning is the most prevalent method. You have input variables (x) and an output variable (Y) in supervised learning, and you use an algorithm to learn the mapping function from the input to the output, i.e. Y = f. (X).

The purpose of supervised learning is to properly estimate the mapping function so that you can forecast the output variables (Y) given fresh input data (x).

Because the process of an algorithm learning from the training dataset may be thought of as a teacher overseeing and monitoring the learning process, it is termed supervised learning. We know the right answers, so the algorithm iterates over the training data, making predictions that are then corrected by the instructor. When the algorithm achieves a desirable level of performance, the learning process comes to an end.



However, there are two sorts of issues that might arise as a consequence of this form of machine learning:

1) Classification Problem

This entails taking the input data and, using the training data, determining which class/category it belongs to. This is used when the projected output is a category, such as red, blue, or green.
An algorithm that ranks automobiles by color or kind, for example.

2) Regression Problem

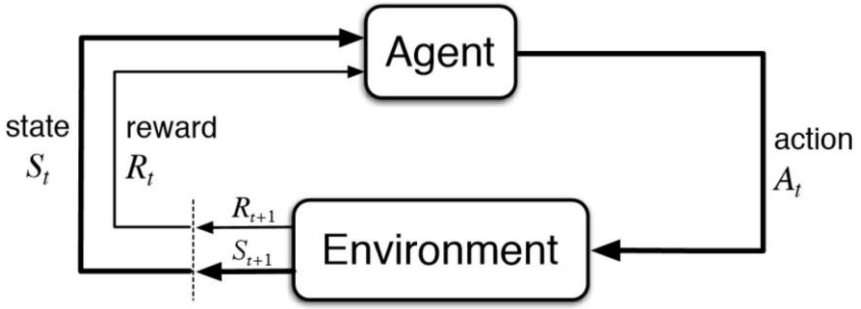
Regression is the process of determining the connection between two or more variables. When the desired result is a real value, such as weight or dollars, it is employed.

For instance, an algorithm that calculates the amount of energy used in a certain period of time for a specific nation

The major reason why many people believe supervised learning isn't as good as it seems is because it's scalability is restricted. Not only that, but getting it to a near-perfect state is quite pricey.

Reinforcement Learning

Reinforcement Learning is a form of machine learning technique that enables a system to learn its behavior by receiving input from its surroundings. This is a behavior that may be learned once and for all, or it can evolve through time. Some Reinforcement Learning algorithms may converge to the global optimum, which is the ideal behavior that maximizes the reward, if the challenge is correctly formulated.



This automatic learning technique eliminates the necessity for a human expert who is familiar with the application area. This kind has the benefit of requiring less time to develop the model since complicated sets of rules do not need to be handcrafted. Not only that, but someone who is skilled in Reinforcement Learning is all that is required.

Computational Programs

Hackystat, Leap, and PSP are just a few of the computational applications that may help with software engineering process measurement. These are the platforms that I will go through in detail.

Hackystat

Hackystat is an open source framework that is used to gather and analyze data from product and software developers automatically. The Hackystat platform's goal is to offer an expandable mechanism that can drastically minimize the overhead involved with collecting a broad range of software engineering data. In addition, the platform has a toolset of analysis that may be used to create a valuable report. Hackystat is frequently utilized in a variety of applications. This may encompass education in the classroom, as well as software engineering for high-performance computer systems.

Leap

The Leap toolkit was developed to give software developer improvement tools that are lightweight, empirical, anti-measurement dysfunction, and portable. Leap allows software developers to collect and analyze personal data such as time, size, errors, trends, and checklists. There are two primary activities on the platform Leap. Gathering raw data and conducting Leap analysis are two of these tasks. Secondary tasks such as improving definitions, checklists, and patterns may help to boost these. Finally, these primary and secondary

tasks may be used to strengthen individual skills or conduct a group evaluation of work results.

PSP

The Personal Software Approach (PSP) is a software development process designed to assist software developers in better understanding and improving their performance by keeping track of their source code's expected and actual development. This platform was built by Watts Humphrey to apply the core concepts of the Software Engineering Institute's (SEI) Capability Maturity Model (CMM) to a single developer's software development processes. The goal of this platform is to provide software engineers with the process skills they'll need to work on a TSP team.

Ethics

Ethics is the study of value notions, whether positive or negative. I'll talk about the ethics of privacy and accuracy in terms of ethics.

Privacy:

Privacy is one of the most important ethical considerations in software engineering. The reason for this is because of unauthorized access to personal information and data abuse.

a) Location Privacy

LTSs (location-based tracking systems) record the whereabouts of things and, in particular, humans using a variety of technologies. Users' security and privacy are harmed when they utilize LTS.

When you use an App that accesses your location, you have no idea what other Apps or corporations have access to your location, which raises the problem of user security and privacy. As a consequence, they may use your information to target you with adverts depending on your location.

For example, with the introduction of mobile applications like Snapchat that monitor a user's position in real time, the adverts on the

app may be tailored to the user's actual location.

a) Public Information

When we use a lot of software or services these days, we have to input our data. Because some data is personal and sensitive, there is a risk of data misuse. This is one of the reasons why IT security is so crucial. However, since it collects our data, it's never quite apparent what they're doing with it.

Personal information, for example, is very valuable to marketing firms. They take user data and utilize it to anticipate people's purchase habits. This aids in the targeting and adaptation of their advertisements. This information may be acquired from various sources to assist businesses with their sales. Although privacy rules exist, in many situations, the user is not required to read them all, and if they do not accept the conditions, they will be unable to access the service. As a consequence, individuals wind up giving their personal information to a corporation for a service without understanding what that firm does with it. There are other corporations, such as Facebook, that update their privacy rules without providing proper warning to its users, for example. Facebook is one of the companies that has done this several time.

Accuracy

Another focus of software engineering ethics is accuracy. Although it is a large subject, it is fraught with ethical concerns.

a) Software Accuracy

The analyst of a system is expected to know and be able to forecast all possible states for complex systems, however this raises a number of ethical concerns about software.

accuracy. Performing the program via system validation and verification is one technique to handle these ethical challenges.

On the surface, it seems that a system developer is morally obligated to remedy all system problems. However, if they deal with faults, it might lead to ethical difficulties since 15-20% of efforts to fix software defects usually result in the introduction of a couple of new errors.

Fixing a mistake in a program with more than 100,000 lines of code, for example, has such a high risk of producing a more serious problem that it is often preferable to leave the fault alone and work around it rather than attempt to repair it.

b) Language and Culture

The vocabulary and language used to ask a question may have a big impact on the accuracy of the information obtained. This is true for any system in which the user is compelled to communicate with software using notions that they are unfamiliar with.

Many superstitions revolve on the numbers "4" and "8" in Chinese culture, for example. Because 4 implies "death" and 8 sounds like "rich," this is the case. As a consequence, if an app asks the user to choose their favorite number from a selection of numbers, 4 will have a very low number and 8 will have a very high number. However, in other nations, similar superstitions may not exist, or other superstitions involving other numbers may exist, causing the statistics to be less accurate than the app's creators would want. As a consequence, understanding diverse cultures and languages is crucial when it comes to data analytics ethics.

References

1. http://web.stanford.edu/class/cs234/index.html
2. http://reinforcementlearning.ai-depot.com
3. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.119.8164&rep=rep1&type=pdf
4. http://ieeexplore.ieee.org/document/5155910/?reload=true
5. https://www.scu.edu/ethics/focus-areas/internet-ethics/resources/unauthorized-transmission-and-use-of-personal-data/
6. https://www.revolvy.com/main/index.php?s=Instruction%20path%20length
7. https://www.researchgate.net/publication/228731372_Experiences_with_hackystat_as_a_service-oriented_architecture
8. https://pdfs.semanticscholar.org/d0c3/5a51a50bd8f8f348a8bb2d0298b5ba83cec0.pdf
9. https://en.wikipedia.org/wiki/Personal_software_process
10. https://www.mathworks.com/discovery/unsupervised-learning.html
11. https://link.springer.com/article/10.1007/BF02249053
12. http://repository.cmu.edu/cgi/viewcontent.cgi?article=1231&context=sei
13. https://en.wikipedia.org/wiki/Source_lines_of_code
14. http://www.projectcodemeter.com/cost_estimation/help/GL_sloc.htm
15. http://www.expertsystem.com/machine-learning-definition/
16. https://careermemo.wordpress.com/tag/data-analysis/
17. https://machinelearningmastery.com/supervised-and-unsupervised-machine-learning-algorithms/
18. https://www.slideshare.net/tonmoybhagawati/presentation-on-supervised-learning
19. https://dev.to/overridedeveloper/an-intro-to-supervised-learning-the-good-the-bad-and-the-ugly-1oo