

ソフトウェアメトリクス ーソフトウェア開発データの解析技法

平成23年度シラバス

2011年2月1日

国立情報学研究所

トップエスイープロジェクト

代表者 本位田 真一

1. 講座名

ソフトウェアメトリクス —ソフトウェア開発データの解析技法

2. 担当者

野中 誠、水野 修、阿萬 裕久

3. 本講座の目的

本講座の目的は、ソフトウェア開発、プロジェクト管理、品質マネジメントに関する各種のアクティビティにおいて、定量データに基づいた意思決定を行うための技法とその活用スキルを習得することである。到達目標は、ソフトウェアプロダクトおよびプロセスに関する基本的なメトリクスの意義と活用方法を理解すること、それらのメトリクスにより得られたデータの解析が行えること、ソフトウェア開発マネジメントにおける局面で適切な解析技法を選択して適用できることの3点である。

4. 本講座のオリジナリティ

ソフトウェア開発のマネジメントには「アート」と「サイエンス」の両面が必要だが、本講座では「サイエンス」の側面に重点を置いている。すなわち、プロジェクト管理や品質マネジメントにおける実践的・経験的なノウハウの伝授よりも、データで事実を把握し、理論モデルや各種の統計解析技法を用いて意思決定に役立つ情報を得ることに焦点を当てている。これを、ソフトウェア開発に関する豊富なデータを用いながら、統計解析ツールを活用して演習主体で学ぶことが本講座の特長である。

5. 本講座で扱う難しさ

ソフトウェア開発マネジメントの実務において、定量データに基づいた意思決定が広く浸透しているとは言い難い。その理由として、そもそも信頼できるデータが収集できていないという問題もあるが、一方で、収集したデータに対して適切な解析技法を適用できていないという問題もある。最近では高度な統計解析技法をパソコンで手軽に利用できる環境が整ってきているが、どのようなソフトウェア開発データに対してどのような解析技法を適用すればよいのか、体系的に整理された情報は少ない。本講座ではこの溝を埋めるべく、データに対して適切な解析技法を適用し、その結果から意思決定に役立つ有益な情報を得るプロセスを重点的に扱う。

6. 本講座で習得する技術

本講座で習得目標とする技術は次の通りである。

- ・ ソフトウェアプロダクトおよびプロセスを定量的に表現する各種メトリクス
- ・ 欠陥数の予測技法（理論モデル、計数データの回帰分析）
- ・ fault-prone モジュールの判別技法（ロジスティック回帰分析、リポジトリマイニング）
- ・ 開発工数の予測技法（理論モデル、多変量回帰分析）
- ・ 信頼性の評価技法（静的モデル、動的モデル）
- ・ プロジェクト成否要因の分析（ロジスティック回帰分析、構造方程式モデリング）
- ・ 各種の可視化技法（テスト管理図、多次元尺度法など）

7. 前提知識

特定の前提知識を必須とはしないが、次の項目を理解したうえで受講することが望ましい。

- ・ 基本的なプロダクトメトリクス（ソースコード行数、サイクロマティック数）
- ・ オブジェクト指向設計の基本概念（クラス、インスタンス、メソッド、メッセージ等）
- ・ ソフトウェア開発ライフサイクルに関する基本知識（主な工程とそのアクティビティ）
- ・ 統計学の基礎（基本統計量、尺度水準、相関と回帰、正規分布、検定）

8. 講義計画

第1週：イントロダクション –ベンチマークデータを用いた議論

第2週：プロダクトメトリクス –コード行数、結合度、凝集度、オブジェクト指向設計メトリクス

第3週：プロセスメトリクス –欠陥除去率（累積、工程別）

第4週：欠陥数の予測(1) –レイリーモデル

第5週：欠陥数の予測(2) –計数データの回帰モデル

第6週：fault-prone モジュールの予測 –判別分析、ロジスティック回帰分析

第7週：開発工数の予測(1) –COCOMO II モデル

第8週：開発工数の予測(2) –重回帰モデル

第9週：信頼性の解析(1) –静的モデル

第10週：信頼性の解析(2) –動的モデル（信頼度成長モデル）

第11週：混乱プロジェクトの判別（水野）

第12週：ソフトウェアリポジトリマイニング（水野）

第13週：品質の可視化 –テスト管理図、多次元尺度法

第14週：品質への影響要因のモデル化 –構造方程式モデリング

第15週：関連トピックス、まとめ

9. 教育効果

データを実際に解析しながら各種技法を学ぶことで、理解を深めることができる。また、演習で用いるデータはいずれもソフトウェア開発に関わるデータであるため、実務での応用に結びつけることができる。

10. 使用ツール

- R：統計解析向けプログラミング言語、及びその開発実行環境。オープンソース。フリーソフトウェア。
- Weka：データマイニングツール。オープンソース。フリーソフトウェア。
- Microsoft Excel：表計算ソフトウェア。

11. 実験及び演習

ノートパソコンに使用ツールをインストールし、実際にデータを解析する。

12. 評価

出席日数とレポート課題への取り組み状況を総合して評価する。

13. 教科書/参考書

教科書は指定しない。参考書は次の通り。

- Linda M. Laird and M. Carol Brennan, 野中誠・鷺崎弘宜訳, 演習で学ぶソフトウェアメトリクスの基礎, 日経 BP 社, 2009.
- Norman E. Fenton and Shari Lawrence Pfleeger, Software Metrics: A Rigorous and Practical Approach, 2nd edition, Course Technology, 1998.
- Stephen H. Kan, 古山恒夫・富野壽訳, ソフトウェア品質工学の尺度とモデル, 構造計画研究所, 2004.
- John McGarry, Cheryl Jones, Elizabeth Clark, David Card and Beth Layman, 古山恒夫・富野壽訳, 実践的ソフトウェア測定, 構造計画研究所, 2004.
- Katrina D. Maxwell, Applied Statistics for Software Managers, Prentice Hall PTR, 2002.
- Roger S. Pressman, 西康晴・榊原彰・内藤裕史訳, 実践ソフトウェアエンジニアリング, 日科技連, 2005.
 - ◇ 15 章 成果物に関するソフトウェアメトリクス
 - ◇ 22 章 プロセスとプロジェクトのメトリクス
 - ◇ 23 章 ソフトウェアプロジェクトの見積り