

AmazonAPI を用いたレビューの信頼度向上

PM コース 矢吹研究室 1242042 齋藤 勇也

1 研究の背景

この研究の背景には2つの理由がある。第1は時代ごとの商品の購入方法の変化生じたこと。第2は現在のレビューの信頼性に疑問を持ったことである。

商品の購入方法は十年前に比べ、大きな変化が生じている。それに伴い、商品のレビューの伝わり方にも変化が生まれていると考えられるだろう。そこで過去と現在の商品の購入方法及びレビューの伝わり方を比べてみた。

過去の商品の購入方法は商品の下に足を運び自身の手で触れていたのが大半であり、商品を購入した人物が知っている特定数の相手のみにレビューを語るといった限られた表現しか行えていない。しかし現在では仮想空間でも商品の売買が行いやすい。このことにより現在は商品に対して不特定多数の人々がレビューを書き込み、それらのレビューをネットワークに接続できるすべての人がレビューを閲覧することが可能である。つまり本人の知らない人物のレビューに対して評価を下さなければいけない環境下になり、より客観的にレビューを判断する必要が出来たと言えるだろう。

また、レビューの評価方法に疑問を持ち始めたことも背景にあげられる。

現在のレビューは数値評価に対して人数を割った平均値を表示している。だが商品とは一切関係のないレビューや商品に対して明らかに商品に対して理解が足りないレビューがあり、それらのような参考にならないレビューも多々存在する。従って参考にならないレビューを排除し、より参考になるレビューのみを利用したいと考えたからだ。

2 研究の目的

レビューに対して重みを付け、平均値に比べて平均値に重みを加えたものがより信頼性があることを立証することを目的とする。

3 プロジェクトマネジメントとの関連

PMBOKにおける10個の知識体系エリアのひとつとして品質マネジメントがある。その内容では「品質コントロールの最終的な目標は、成果物の正しさを決定することである。」(プロジェクトマネジメント知識体系ガイド(PMBOKガイド)第5版)と記述されている。この研究によりレビュー評価の信頼性が向上することで成果物の正しさをより証明できるものとなるだろう。

4 研究の方法

4.1 計算方法の策定

計算方法は商品一件あたりに存在するすべてのレビュアーの数、レビューごとの参考になったと答えた比率から割り出すこととした。

具体的には

$$y = \frac{5 \text{ 段階評価} * \frac{\text{「参考になったと回答した人数」}}{\text{参考になる・ならないの判断をした人数}}}{\frac{\text{「参考になったと回答した人数」}}{\text{参考になる・ならないの判断をした人数}}}$$

このような計算方法を使用した。

4.2 計算方法の策定

amazon にはレビューを見た人に対して参考になったかどうかを判断させるシステムが存在するのでこれを利用することとした。

amazon のレビューシステムは 5 段階評価であり、これらの評価値をレビュアーが主観的に判断し、その理由をコメントできるシステムである。

これら amazon のレビューデータを集める際に、2 つの方法でデータを収集した。

1 つめがエクセルを利用し、データを手入力で行い計算式を入れていく方法である。この方法はエクセルは補助的役割しか果たしておらず、ほぼ手作業なのでミスが生じやすい。よってある程度データが取れ次第 2 つめの方法を行うこととする。

2 つめが Linux のコミュニティにより開発されているオペレーティングシステムの Ubuntu の環境を準備し、amazonAPI を利用し必要なデータのみを受信、プログラムに計算を行わせる方法であるこの方法は amazonAPI を利用することで 1 つの商品のレビューすべてを一度に受け取ることが可能になり、これにより短時間で多量のデータを取得することが出来る。

4.3 参考文献

- 雑誌論文 [1]
- 書籍 [2]

5 現在の進捗状況

計 86 件のレビューデータを収集し、そのデータを下に平均五段階評価と重み付き 5 段階評価の散布図を作成した。

6 今後の計画

評価段階ごとの参考になった人数を記録し、クラスター分析や標準偏差を求めさらに分析を進める。

参考文献

- [1] 山澤美由起, 吉村宏樹, 増市博. Amazon レビュー文の有用性判別実験 (解析, 分析), 2006.
- [2] プロジェクトマネジメント知識体系ガイド (PMBOK ガイド), 第 5 巻, 2013.