

タイトル：ロジスティック回帰を用いたMCバトルにおけるMC推定（言語処理学会 第28回年次大会 発表論文集より）

著者一覧：吉實隼稀、但馬康宏（岡山県立大学大学院 情報系工学研究科システム工学専攻）

まとめ：本研究はロジスティック回帰を用いてMCバトルの歌詞を解析し、各MCの個性を捉えることが目的である。

主張：地名を発した数やアンサー単語の平均回数の対戦相手との差が、MCの個性を推し量るに有効な説明変数であることが示され、MCごとの個性がベクトルの角度により捉えられた。

興味深いと思った点：もともとMCバトルを見るのが好きであるため、この論文を選んだ。また、MCバトルの歌詞を解析するという手法を使う以上、フロー重視のラッパーや、ダブルミーニングなどがどのように評価されるのか気になった。実際、フローは対象外とされており、ニガリはこれが原因で低い評価になったとのことだが、それならば、mu-

tonの評価が高くなったのが意外だった。ディスも対象外になっているようで、これにより呂布カルマの評価が低くなったとした。罵倒や、言葉に込められた感情の研究からヒントが得られそうだと感じた。パンチラインスコアはTF-

IDFを用いて算出しているとのことだが、人間がどのくらい感情を揺さぶられるかといった要素を、どのようにして計算しているのだろうと疑問に思った。実際に著者も、各MCから受けた印象と、実際のパンチラインスコアの間にギャップを感じているようである。R-

指定もそうだが、呂布カルマのパンチラインスコアが低くなっているあたり、改善の余地がありそうである。ヴァイヴス単語として「だ」「ぜ」「たい」「だぜ」「ねえ」「だろ」

「だぞ」「だよ」「じゃねえか」「ねえよ」が挙げられているが、これがすべてヴァイヴスとして勘定してよいのか、またこれ以外には存在しないのか気になった。

ライムも脚韻だけを対象としているが、母音が5つである日本語で、頭韻や、文中に出てくる韻などを検出する方法がないだろうかと思った。KOKという大会では、MCのスキルをいくつかの項目に分け、格闘技のように、9対10などとして数値にして審査していると聞いたので、このような研究が進めば、MCバトルの大会で実用化することもできるかもしれない。

次に読むべきもの：pythonライブラリであるscikit-

learnのロジスティック回帰の手法について書かれたもの。また、ランダムフォレストを用いた研究も行われている。他にも、SVM、DNNなど。感情分析の研究とも関わりがあるようだ。TF-

IDFについて書かれたもの。地名の抽出に使われた形態素解析のjanomeについて書かれたもの。歌詞や

テキストデータの分析、分類を行った研究は多数存在しているらしい。品詞に着目した研究もある（そういえば、歌詞に出てくる言葉をもとに、ツリーマップのようなものを作ったサイトも見たことがある）。