

# 中間発表

04T4027T

斉藤 章久

# 概要

- インターネットが普及に伴い、インターネット上の掲示板の利用者の数も増えています。

利用者は自分の興味がある分野の掲示板にアクセスし、知りたい情報を手に入れたり、趣味の合う人同士で交流をしたりすることができます。

- しかし、この掲示板で知りたくない情報を見ってしまうことがあります。

そういう情報の中の一つに**ネタバレ**というものがあります。

# ネタバレとは？

ネタバレは、まだ自分が見ていない漫画や映画の内容を先に掲示板に書かれてしまうことです。

来週の内容を楽しみにしている人にとってはとても苦痛です。

このネタバレの文を検出しようというのが本研究となります。

# ネタバレ例(1)

緑間「くっ...」 かがみんに抜かれる。宮地（一体なんなんだよコイツは...!?) 木村（ついこの間まで中坊だった奴に ウチが押されてるってのかよ!?)

大坪（全国でも見たことがない... なんだこの...常軌を逸した跳躍は...!?) 大坪木村を越えてダンクを決めるかがみん。

「うわああ高ええー!!」「一人で秀徳を圧倒してるぞ!?’小金井「スゲーなニス火神!!」 かがみん「もっとガンガンボールくんねーですか」「オレー人で決めてやる...!!」

秀徳監督（日本人ばなれした信じられんバネだ... これをくい止めるのは...容易ではないな...）残2:01 誠凜45-56秀徳

秀徳ベンチ「監督何か手をうたないと...」 監督「んー...そうなんだけど...でもそうするとウチもかなりのリスクが出るんだよなー」

監督「だから今はあえてこのままいく」ベンチ「えっ」 監督「まあ。。あわてるな」「時間がたてばじきに 詰みだ」

# ネタバレ例(2)

来週はヒメコのこと好きなやつの話

二人が本当は付き合ってるのかどうか確かめるために部活見学？の名目で観察するが

ユウスケとヒメコが仲が良いので夫婦か！とか初恋か！といちいちつつ込み。

結局スイッチにこいつらの(自分も含めて)は恋愛とかではなく絆がどうのこうのって言われる

最終的にスケット団の絆にはかなわないけど恋愛は別ですよ  
ねってことでまだ諦めない模様。

かいつまむとこんな感じ

# ロジスティック回帰分析(1)

- 発生確率を予測する手法
- 予測結果はすべて0から1の間の数字
- 従属変数に2値の質的変数を用いる  
(例) Yes or No
- この従属変数の値を実績値として用い、説明変数を用いて発生確率を説明する。

# ロジスティック回帰分析(2)

- 説明変数を $ax+b$ のような形にする。
- 分析結果を $y$ とするが、 $y$ は0から1という範囲があるので、 $y=ax+b$ と結びつけできない場合が出てくる

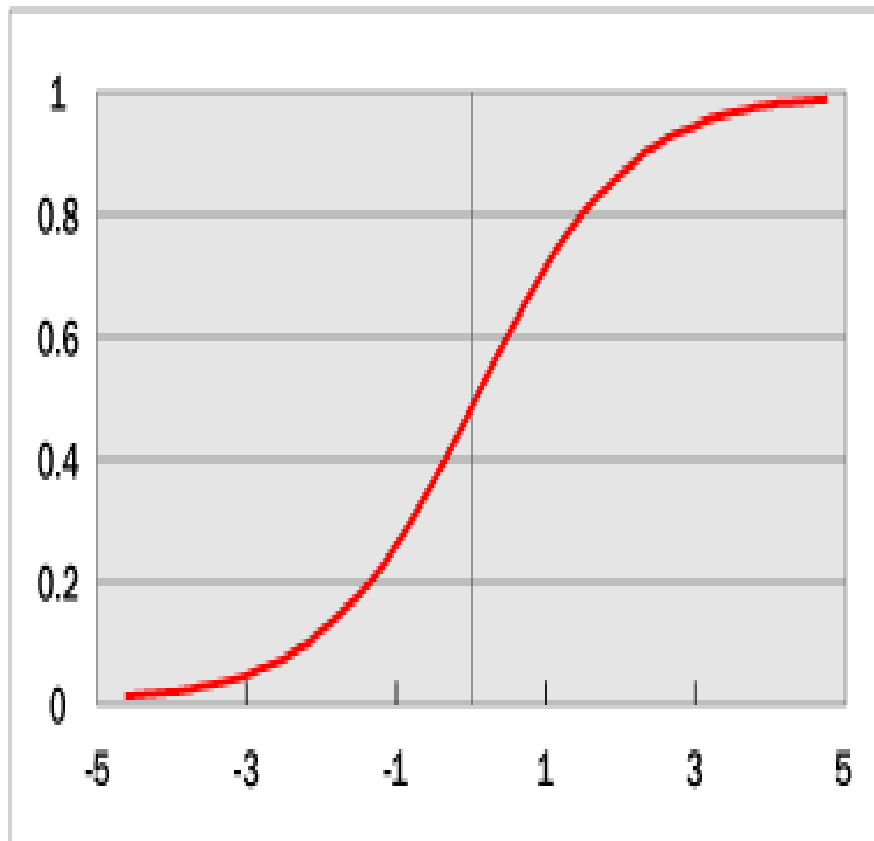
↓ よって

$$\log_e\{y/(1-y)\} = ax+b$$

このような式を用意する。

- $y$ は購入する確率、 $1-y$ は購入しない確率
- これらの $y/1-y$ のことを「オッズ」と呼ぶ

# ロジスティック回帰分析(3)



- この図の横軸は $\log_e\{y/(1-y)\}$ であり、 $ax+b$ である。
- 縦軸は発生確率の $y$ である。

# ロジスティック回帰分析(4)

- この図は $ax+b$ がいかなる値もとりに、かつ $y$ が0から1の間に収束する。
- また、このようなS字を崩したようなデータの形はいろいろな例で使える  
(例)細菌の繁殖の発生から繁殖による隆盛。人間の学習線。

# これからの課題

## ①ネタバレの文の特徴を発見する。

より精度を上げるためには、多くの特徴を発見することが必要なので。

## ②ロジスティック回帰分析の理解度を深める。

まだ調べ始めたばかりで、「オッズ」や「ロジット変換」など曖昧になっている部分があるので、参考文献を読み理解力を深めようと思う。