

ドラマ視聴率の予想

2年 柴山りな

概要

- 朝の連続テレビ小説過去10年分20作品についての情報をインターネット上で集め、ニューラルネットワークで学習をした。来月2018年10月から始まる「まんぷく」の平均視聴率、およびその次の「青空」の平均視聴率を予測した。
- 予測変数として、時代設定と前回の平均視聴率を採用した。
- 結果、「まんぷく」「なつぞら」ともに「半分、青い。」とほぼ変わらない20.6%という予測になった。



「半分、青い。」



「まんぷく」



「なつぞら」

作品を判断する方法

放送時期

- 放送時期が上半期か下半期か

時代設定

- 何年に物語が始まるか

舞台

- どこの地方が舞台になっているか（東北、関東などで分けた）

実在の人物をモデルにしているかどうか

- 主人公が実在の人物をモデルにしているか、脚本家のオリジナルストーリーなのか

主演女優の経験値

- ヒロイン女優が今までに連続ドラマ・映画で何回主演を経験したことがあったか

前回の平均視聴率

- ひとつ前のドラマがどれだけの平均視聴率であったか

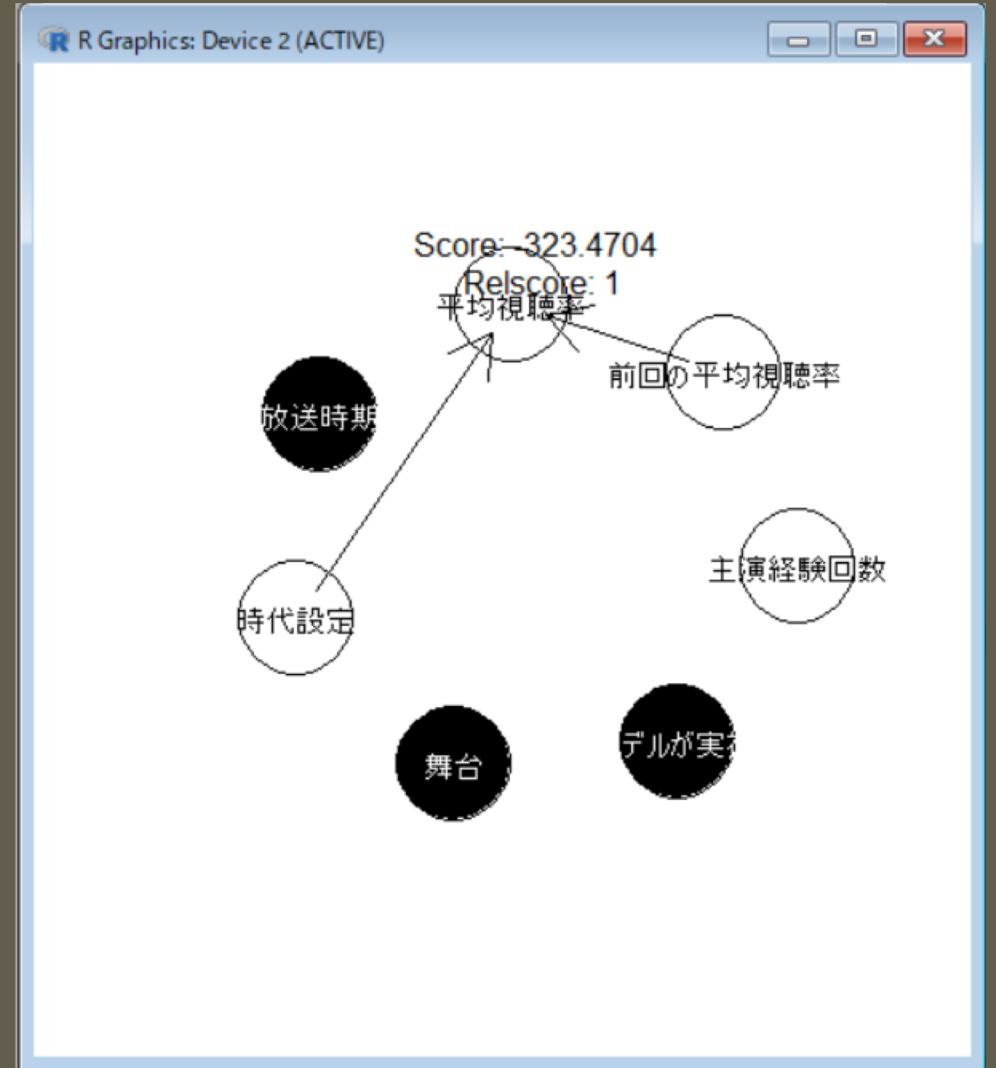
作品を判断する 方法

平均視聴率の原因として考えられる項目6つ挙げ、次のような表を作った。情報はインターネットにあるものを参考にし、それでも時代設定などの不確実な部分については推測で作成した。

	平均視聴率	放送時期	時代設定	舞台	モデルが実在	主演経験回数	前回の平均視聴率
半分、青い。	20.8	上半期	1971	中部	実在しない	2	20.1
わろてんか	20.1	下半期	1902	近畿	実在する	1	20.4
ひよっこ	20.4	上半期	1964	関東	実在しない	8	20.3
べっぴんさん	20.3	下半期	1945	近畿	実在する	2	22.8
とと姉ちゃん	22.8	上半期	1930	関東	実在する	0	23.5
あさがきた	23.5	下半期	1857	近畿	実在する	3	19.4
まれ	19.4	上半期	1994	関東	実在しない	3	21.1
マッサン	21.1	下半期	1920	近畿	実在する	0	22.6
花子とアン	22.6	上半期	1900	中部	実在する	4	22.4
ごちそうさん	22.4	下半期	1911	近畿	実在しない	2	20.6
あまちゃん	20.6	上半期	2008	東北	実在しない	1	17.1
純と愛	17.1	下半期	2000	沖縄	実在しない	0	20.7
梅ちゃん先生	20.7	上半期	1945	関東	実在しない	13	19.1
カーネーション	19.1	下半期	1924	近畿	実在する	7	18.8
おひさま	18.8	上半期	1932	中部	実在しない	7	17.2
てっぱん	17.2	下半期	2008	中国	実在しない	0	18.6
ゲゲゲの女房	18.6	上半期	1939	関東	実在する	6	13.5
ウェルかめ	13.5	下半期	2005	四国	実在しない	0	13.8
つばさ	13.8	上半期	2008	関東	実在しない	4	16.2
だんだん	16.2	下半期	2000	近畿	実在しない	0	15.2

視聴率に影響を与える要因の選択

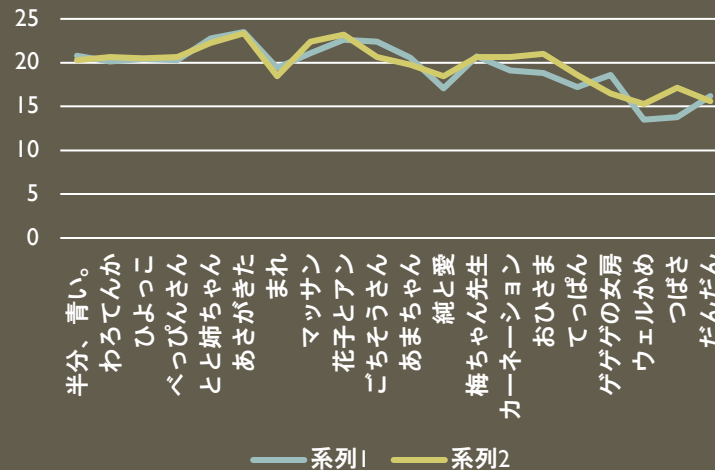
- この6つのうち、平均視聴率に影響を与えている要因を選んだ。
- 選び方は、6つの要因に平均視聴率を加えた7変数で、ベイジアンネットワークモデルを作成した。白丸が連続量、黒丸が離散量を表している。
- ベイジアンネットワークとは確率変数間の依存関係を非循環有向グラフで表現するモデルで、結果は右のようになった。
- ここから推測すると、時代設定と前回の平均視聴率が平均視聴率に影響してる。



来期の視聴率の予測

- 予測変数を時代設定と前回の視聴率、基準変数を平均視聴率として、ニューラルネットワークで学習を行った。
- 学習の結果にデータを適応すると下の図のようになった。
- 水色の線が実際の視聴率、黄緑の線が学習結果を表している。

実際の視聴率と学習結果との比較



来期の視聴率の予測

- 学習結果を適応すると...
 - 「まんぷく」の時代設定(1940)、前回の平均視聴率(20.8)を適応すると
 - 20.61288%
 - 「なつぞら」の時代設定(1946)、前回の平均視聴率(20.6)を適応すると
 - 20.60539%
 - という結果になった。
-
- 時代設定、前回の平均視聴率ともに近い値なので結果もほぼ同じになった。

結果

- 「まんぷく」とその次の「青空」ともにここ2, 3年でほぼ横ばいの20.6%になるという予測になった。

まとめ

- 残念ながら、爆発的なヒットは予測されませんでした。

平均視聴率の推移

