

# 絶対に乗ってはいけないJR

～遅延しない電車にのりたいですよね？～

総合数理学部 2年 渡辺雄斗

# 背景

- 東京都心のJRの列車は、ほぼ毎日遅延が発生している、と聞いていくくらい遅延している。
  - 原因は人身事故や異音の発生など様々あるが、突発的な出来事に対して安全を確実にするためには遅延せざるを得ないため

**「遅延を予測して遅延しない列車に乗ることはできないものか？」**

# 目的

- 遅延を予測といっても突発的な事故は予測できない、、、

JRの様々な路線の過去の遅延結果を分析



遅延の規則性を発見する

# 方法

- 基本的に決定木を用いた分類を行う
- ①**曜日**による遅延発生の方の違を分類
- ②**天気**による遅延発生の方の違を分類
- ③それぞれの予測精度を検証
- なお、遅延情報を利用する路線は
  - 埼京線 ・ 山手線 ・ 総武線 ・ 中央線 ・ 東海道線 ・ 京葉線 の6種類。
- これらのデータを朝（始発～10時）、昼（10時～16時）、夜（16時～終電）に分けて用いる。

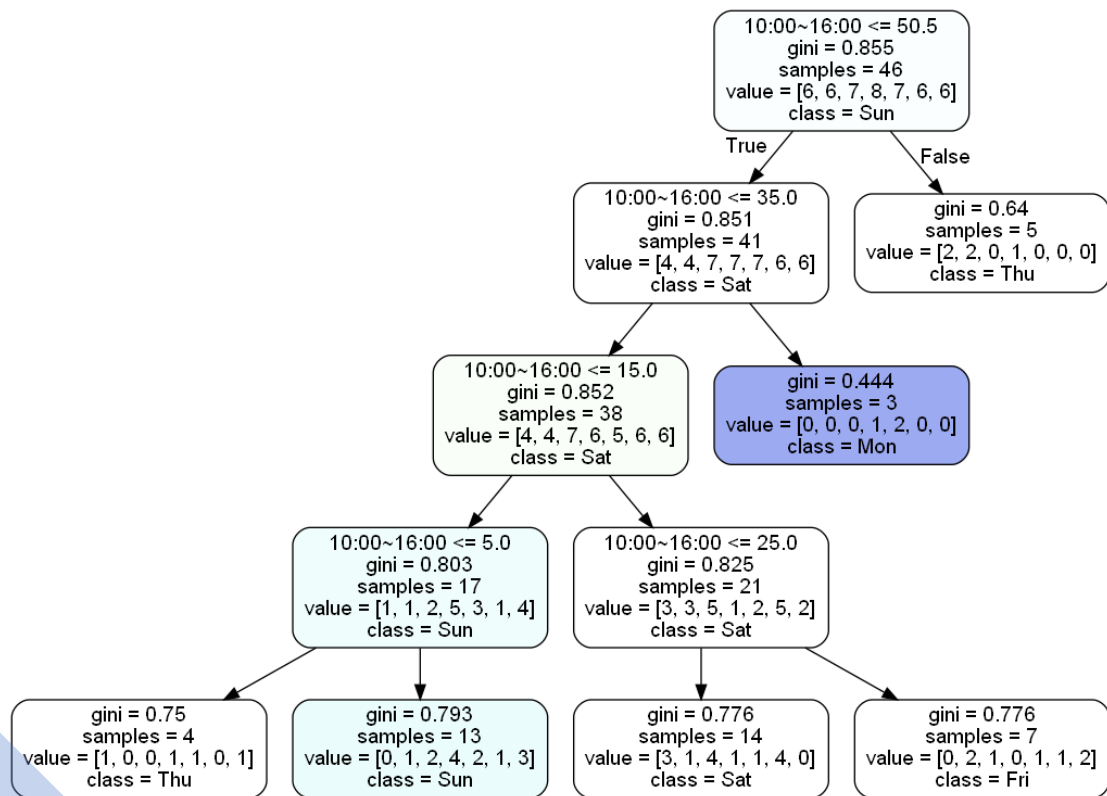
# データ

date	~7:00	7:00~10:00	10:00~16:00	16:00~21:00	21:00~	day	Weather_morning	Weather_lunch	Weather_afternoon
2022/1/12	10	10	10	0	60	Wed	晴	晴	晴
2022/1/11	0	10	0	10	10	Tue	雨	雨	雨
2022/1/10	0	0	0	0	10	Mon	曇	曇	曇
2022/1/9	0	10	10	0	10	Sun	晴	晴	晴
2022/1/8	0	0	0	0	0	Sat	晴	晴	晴
2022/1/7	0	20	10	0	10	Fri	晴	晴	晴
2022/1/6	0	0	0	10	20	Thu	曇	雪	雪
2022/1/5	0	0	0	0	0	Wed	晴	晴	晴
2022/1/4	0	0	0	10	0	Tue	晴	晴	晴
2022/1/3	0	0	0	10	0	Mon	晴	晴	晴
2022/1/2	0	0	0	10	0	Sun	曇	晴	晴
2022/1/1	20	20	20	0	10	Sat	晴	晴	晴

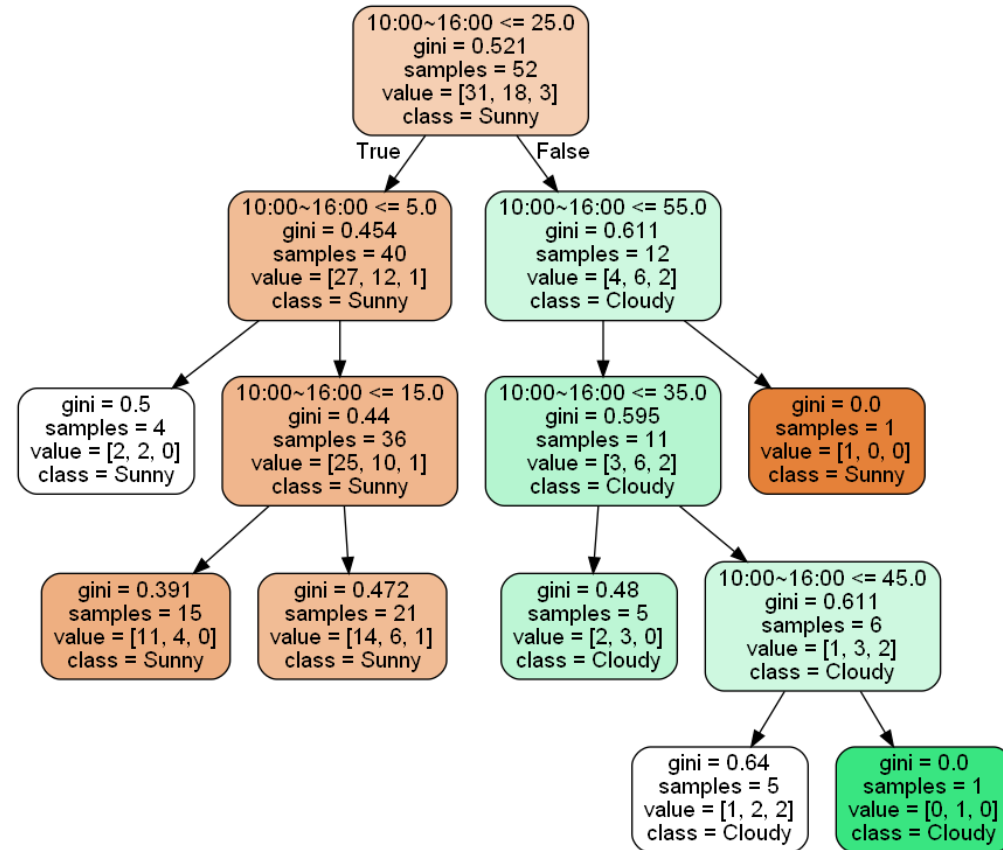
- サンプル数は58個
- ~7:00,7:00~10:00が朝、10:00~16:00が昼、16:00~21:00,21:00~が夜

# 見方の説明

## ①曜日による分類



## ②天気による分類




- 基本的には左に行くほど遅延が少なく、右に行くほど遅延が多い
- true = 以内、 false = 以上
- valueの値 = [0,1,2,4,2,1,3] → [木、金、土、日、月、火、水] に対応する個数



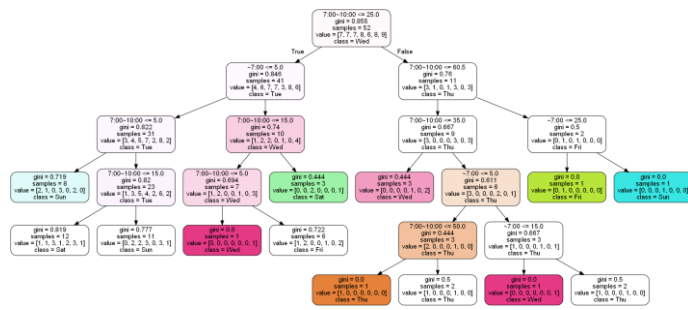
時間がないので分類した結果から面白い結果を得られたものを抜粋

ファイルの中にすべての結果の画像がありますのでみたい人はそちらを見てみてください。

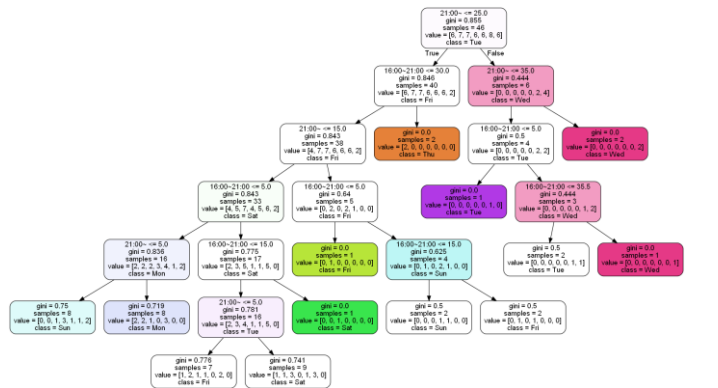


# ① 曜日による分類

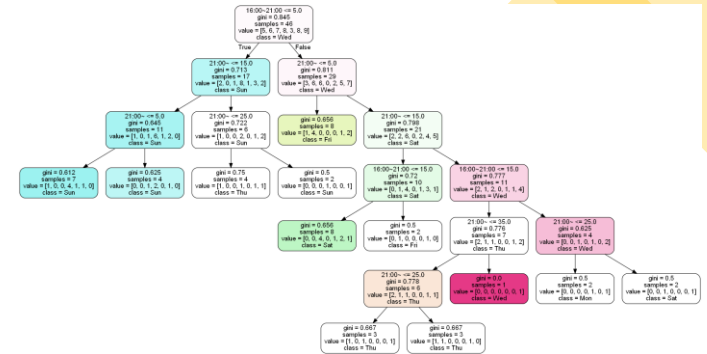
・ 埼京線



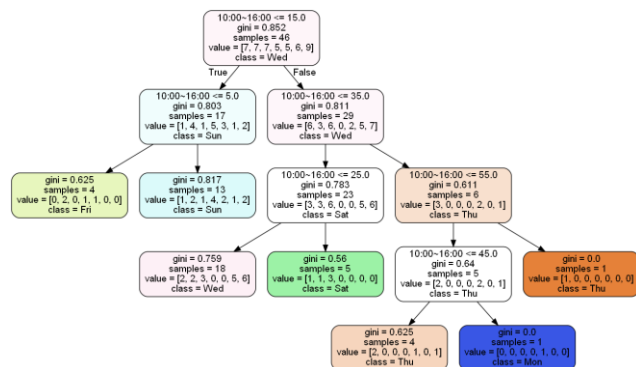
・ 山手線



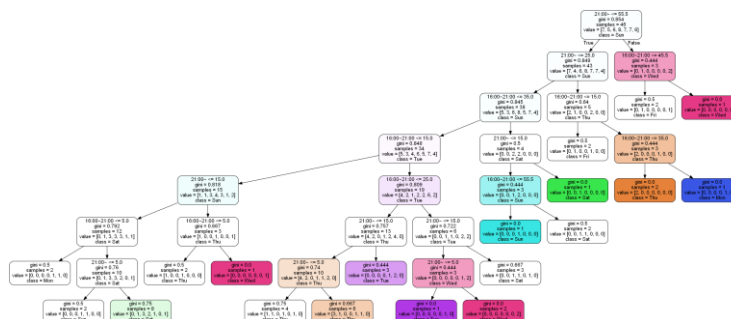
・ 総武線



・ 中央線



・ 東海道線

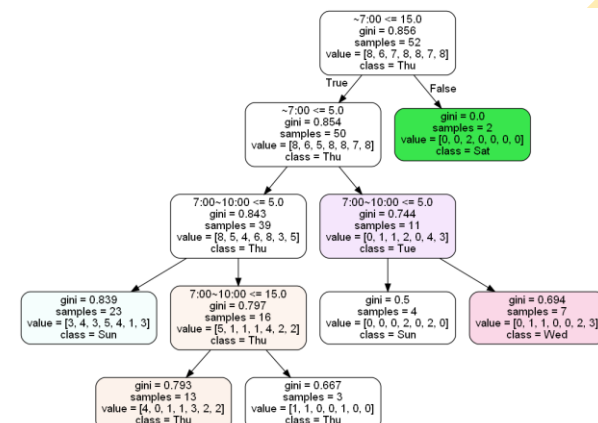


多くの結果は平日に遅延が多く、休日は遅延が少ない路線が多い

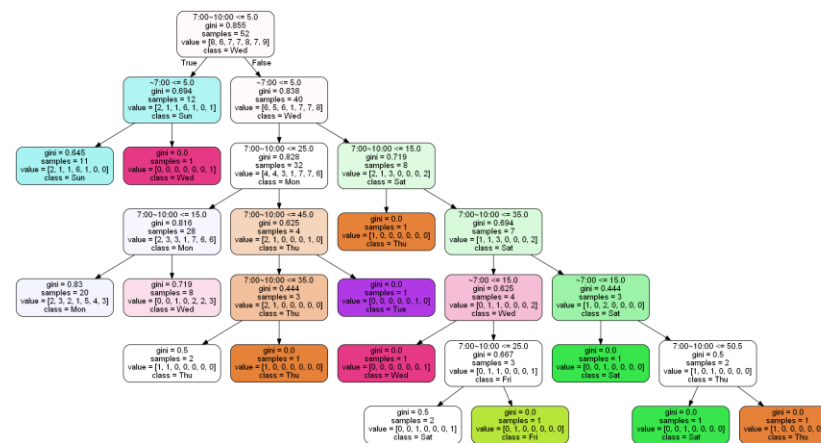


# ① 曜日による分類

- 山手線 午前 (始発～7時～10時)
- 土曜日以外大きな遅延はほぼなし
- 土曜日の遅延も工事などの影響で出てしまった計画遅延の可能性  
→曜日による遅延の違いはさほどない



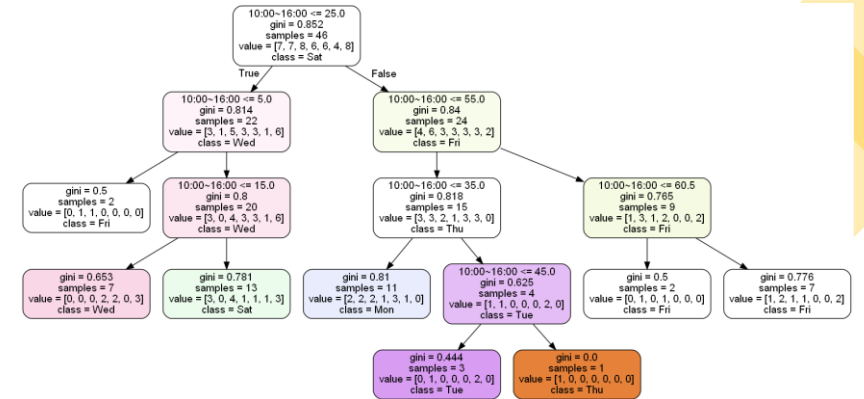
- 中央線 午前 (始発～7時～10時)
- 土曜日に遅延多め
- サンプル数から遅延自体が遅延してない場合の3倍  
→そもそも遅延が多い路線であることがわかる



# ①曜日による分類

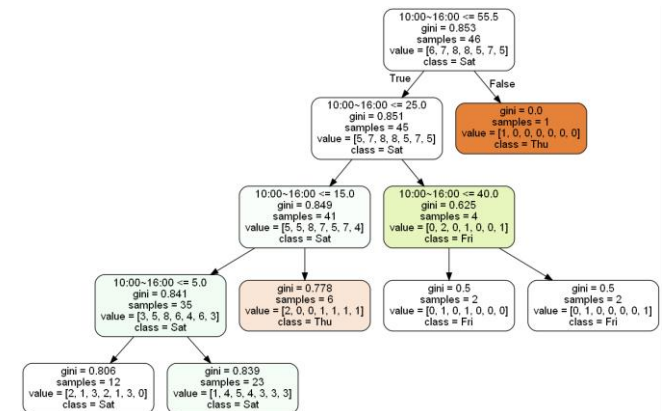
## • 東海道線 昼（10時～16時）

- 金曜の遅延が多い
- 土、日、月の遅延は少ない印象  
→金曜日に少し家の遠めの人が帰るから？



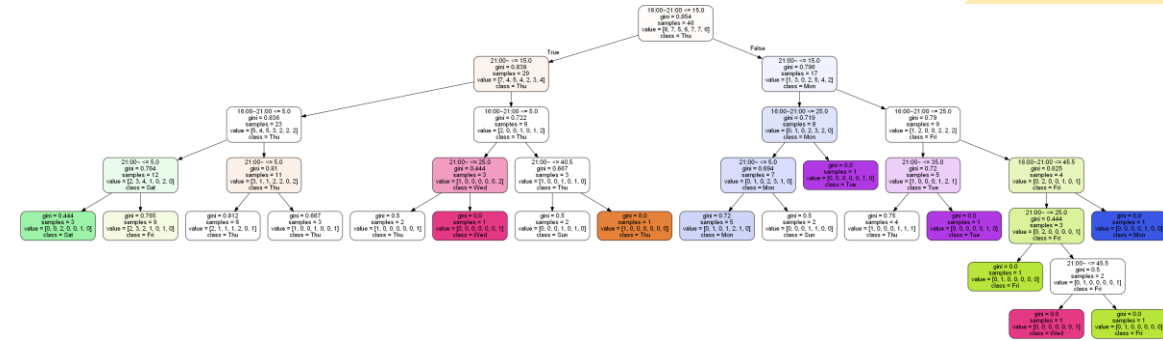
## • 京葉線 昼（10時～16時）

- 左の山に集中している  
→遅延数が圧倒的に少ないことがわかる



# ①曜日による分類

- 京葉線 夜（16時～終電）
- 遅延のある曜日が休日と平日に顕著に分かれている
- 特に土曜日に遅延が少ない  
→平日はやはり夜でも遅延が多いことがわかる

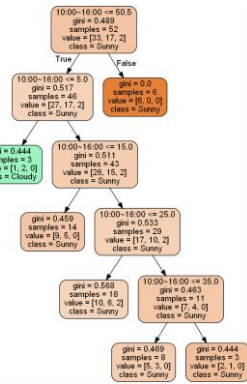


## ①曜日による分類 結果

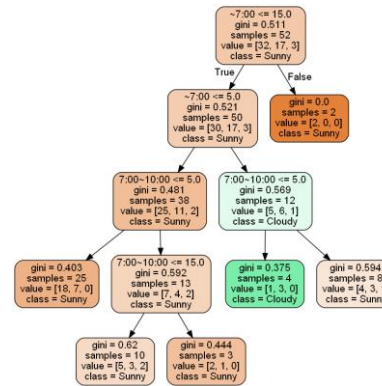
- 基本的に平日の方が遅延が多い
- 特に通勤時間帯に遅延が起こるのが月曜日、水曜日となっている

# ② 天気による分類

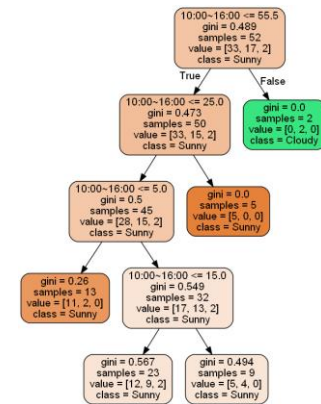
・ 埼京線



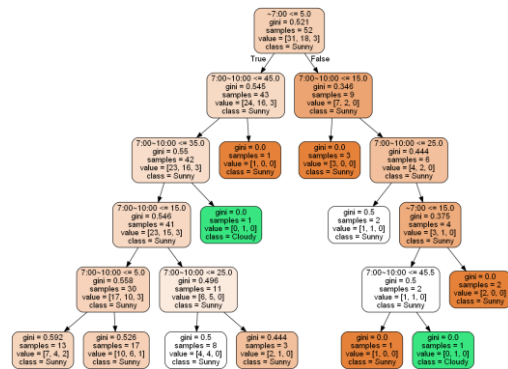
・ 山手線



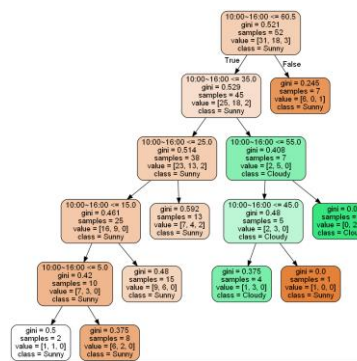
・ 総武線



・ 中央線



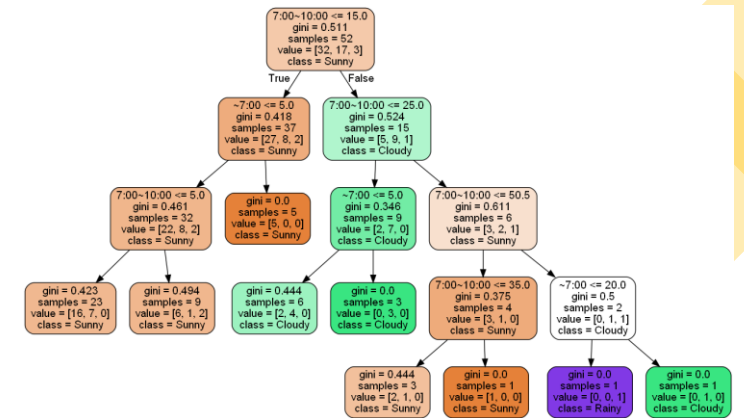
・ 東海道線



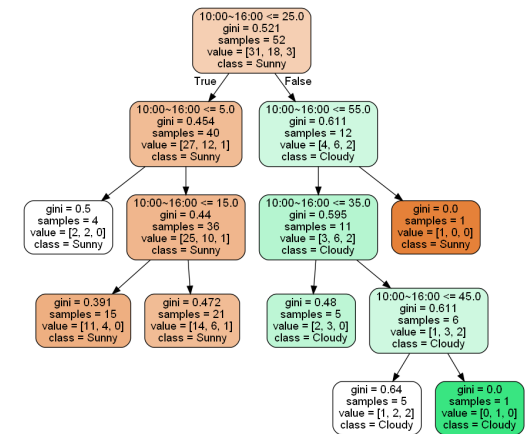
圧倒的に晴れの時のデータが多い  
(オレンジノードは晴)

## ②天気による分類

- 京葉線 午前（始発～7時～10時）
- 曇りの時ばかり遅延が発生している  
→天気による遅延の発生の違いが割と顕著に表れている  
→海辺を走ることから曇りの時は風の影響？



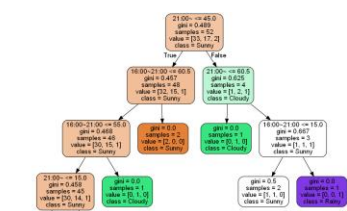
- 中央線 昼（10時～16時）
- 曇りの時の遅延が大きい  
→曇りの日は遅延が大きくなる傾向



## ②天気による分類

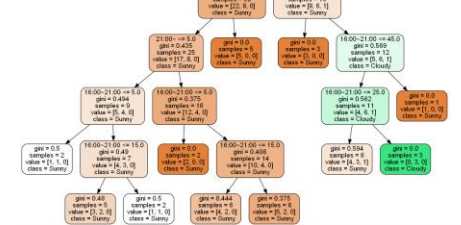
- 山手線 夜 (16時～終電)
- 雨の日の遅延が最も大きい
- 雨に続いて曇りの日の遅延が大きい  
→天気が悪いと遅延が大きくなる

- 埼京線 夜 (16時～終電)
- 山手線同様に天気が悪いときに遅延が大きい



## ②天気による分類 結果

- 天気が悪いと遅延が発生する関係性が見える
- 特に、雨の方が曇りの時よりも遅延が大きくなっている



### ③ 予測精度の検証

今回の検証において、決定木の深さはグリッドサーチを用いて決定した。なお、曜日による分類においては5分割交差検証を、天気による違い分類においては10分割交差検証を行った。その結果、予測精度は下記ようになった。

①曜日による分類 - 学習データ予測精度

	山手線	総武線	京葉線	埼京線	中央線	東海道線
朝	0.32692308	0.38461538	0.42307692	0.40384615	0.48076923	0.42307692
昼	0.26086957	0.30434783	0.2826087	0.32608696	0.32608696	0.36956522
夜	0.47826087	0.47826087	0.47826087	0.58695652	0.60869565	0.56521739

②天気による分類 - 学習用データ予測精度

	山手線	総武線	京葉線	埼京線	中央線	東海道線
朝	0.38461538	0.48076923	0.53846154	0.51923077	0.46153846	0.55769231
昼	0.34615385	0.44230769	0.42307692	0.42307692	0.34615385	0.48076923
夜	0.55769231	0.55769231	0.61538462	0.67307692	0.65384615	0.69230769

①曜日による分類 - テスト用データ予測精度

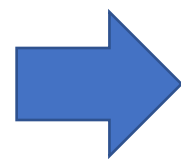
	山手線	総武線	京葉線	埼京線	中央線	東海道線
朝	0.16666667	0.33333333	0.16666667	0.16666667	0.16666667	0.16666667
昼	0.08333333	0.08333333	0.08333333	0.08333333	0.33333333	0.08333333
夜	0.08333333	0.08333333	0.25	0.08333333	0.08333333	0.08333333

②天気による分類 - テスト用データ予測精度

	山手線	総武線	京葉線	埼京線	中央線	東海道線
朝	0.33333333	0.16666667	0.16666667	0.16666667	0.16666667	0.33333333
昼	0.5	0.33333333	0	0.16666667	0.33333333	0.33333333
夜	0.16666667	0.16666667	0	0	0	0.16666667

# 結果考察

- 遅延しない電車に乗るためには、、、
- 平日に乗らない
- とくに月曜日と水曜日の早朝
- 曇りの日は乗らない
- 雨は論外



つまり、  
晴れている休日に乗る

これらのことから、天気と曜日を見ることによって遅延しているかどうかを予測することは可能ではあるということがわかる



# 感想

- 思ったよりも天気と遅延の関係性がはっきりしていて驚いた
- もう少し良い予測方法があったのではないかと思ったが自分の頭では答えに行きつかなかった

# 参照

機械学習実践 (ハイパーパラメータ調整)

[https://free.kikagaku.ai/tutorial/basic\\_of\\_machine\\_learning/learn/machine\\_learning\\_hyperparameters](https://free.kikagaku.ai/tutorial/basic_of_machine_learning/learn/machine_learning_hyperparameters)

JR東日本

<https://www.jreast.co.jp/>

東京の過去の天気

[weather.goo.ne.jp](http://weather.goo.ne.jp)