



1. 本研究について
2. MySQL

経済に関するオルタナティブ・データを考慮した金融マーケット予測手法の開発

Development of Financial Market Forecasting Method
Considering Alternative Data on Economy

Itaru Aso

Graduate School of Information Engineering, Toyama Prefectural University
t855001@st.pu-toyama.ac.jp

Toyama Prefectural Univ.



1.1 取り組み

取り組み内容

1. 本研究について
2. MySQL

- ① MySQL にデータを格納する上で発生した問題を解決
- ② Python から MySQL に格納してあるデータを抽出



2.1 MySQL

1. 本研究について
2. MySQL

発生した問題

重複した内容のツイートが何件か格納されていた

原因

Twitter API からデータを取得するにあたり時期が被ってしまっていたため同じ内容のデータを取得してしまっていた

そのため、重複した内容のデータを削除する必要がある



2.2 問題解決への取り組み

1. 本研究について

2. MySQL

	id	date	week	tweet
	0	2019-01-07 11:43:09	Monday	チームDのMC-U。 カーディガンの赤は血染め。 顔のシミはタトゥー。
	1	2019-01-06 14:43:49	Sunday	まず、お等の練習しろ。
	2	2019-01-06 03:38:58	Sunday	本日ラジオ。 新年一発目は恒例の DDUメンバーが登場。
	3	2019-01-01 14:12:03	Tuesday	あけー。 ことでーす！
	4	2018-12-31 07:56:34	Monday	良いお年を。
	5	2018-12-30 07:31:22	Sunday	本日ラジオ。 年末ラストは歌合戦。
	6	2018-12-28 06:54:00	Friday	
	7	2018-12-25 16:01:53	Tuesday	あと一回だけスプラトゥーンやって寝よう。。。
	8	2018-12-25 16:00:09	Tuesday	ねるー！
	9	2018-12-25 15:51:49	Tuesday	全員寝る！ おまえはいつも悲壮感を漂わせるな。

Figure: MySQL への挿入

`id` を格納するカラムを用意することによって, `id` を利用して重複したレコードを削除する

処理内容

重複したレコードを抽出し, `id` の小さい方のレコードを残し, それ以外のレコードは削除する



2.3 MySQL からツイート取得

ツイート取得

SQL 文の SELECT を用いてデータを取得

1. 本研究について
2. MySQL