

日経平均データについて

日経平均データは1時間足、30分足、5分足のデータが3年分集まった。

Twitter データ収集について

SNS を用いた株価予測として、Twitter を利用する。

使用するの、SNS として発信されたデータ数のみとする。日経平均の株価予想をするため、日経平均の構成銘柄をキーワードとして Twitter のデータを取得する。1日に発信されたデータ数を収集し、ニューラルネットワークに学習させる。現在、Twitter よりデータを取得するプログラムは完成し、日経平均 225 社中 25 社についてはデータ収集を開始もしくは終了している。8月頭までには 225 社すべてのデータ収集が完了できる。

ネットワークに関して

SNSデータがまだ十分に集まっていないので、SNSの データを使わないニューラルネットを作った。時刻 t までの値から時刻 $t+1$ の値を予測するリカレントニューラルネットワークを作成した。周期性のある データに対しては概ね良さそうな出力を得られたが、移動平均乖離率のような不規則なデータに対しては一定の値を返すように学習してしまいうまくいかない。

まとめ

データ収集は日経平均は集め終わり、SNS のデータの収集率は目標の 15%ほどである。ネットワークは現在構築中であり、シミュレータもまた現在製作中である。

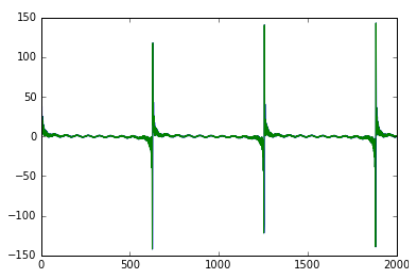


図 1 入力

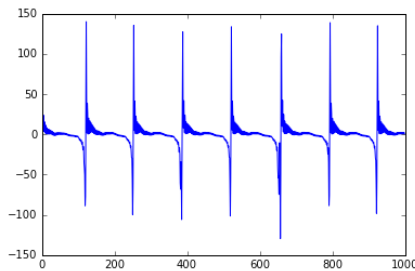


図 2 出力

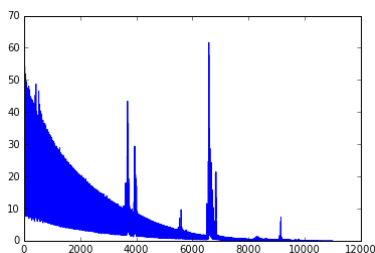


図 3 誤差関数