

情報システム工学科 平成 20 年度後期「自主课题研究」
研究テーマ：サポートベクター回帰を用いた天気予報
学年：3 名列番号：039 氏名：西部匡洋

1. 研究目的

今回の研究では統計解析ソフトウェア R を用いて統計的な観点から天気の予測を試みた。そして予測の正答率がどのくらいになるかについて調べた。

2. 研究内容・方法

天気の予測を行う流れは以下のようになる。まず過去の気象データを統計処理ソフトウェア R に読み込ませる。次に得られたデータをサポートベクター回帰によって学習させることで回帰曲線を得る。そして得られた回帰曲線から天気を予測する。このようにして行う天気の予測を様々なデータの選び方でいい、予測の精度の変化を調べる。

3. 実験に用いたデータ

回帰を行うためのデータとして Yahoo! 天気情報 (<http://weather.yahoo.co.jp/>) から全国各地の過去の気象データを用いた。データの取得は wget というプログラムを使って指定した Web ページをダウンロードし、Perl スクリプトで必要な情報を Web ページから抜き出すことで得られる。

4. 回帰方法

回帰予測にはサポートベクター回帰を用いた。パッケージは

kernlab、関数は ksvm、カーネルは多項式カーネルである polydot を用いた。

5. データの選び方

- ・前日予測を行う際に過去何日分のデータを用いるか？
- ・何日前予測か？
- ・対象地域による違いは？(東北、関東、北陸、九州)などの違いに注目した。

6. 結果・考察

結果を下図に示す。図より正答率は 5 割～7 割の間を推移していることが分かる。また金沢以外の 3 地域では過去 1～2 日分のデータを用いたときに一番精度がよく、それ以降はほぼ右肩下がりになっていることがわかる。何日前に予測するかの結果については各々の地域で精度にばらつきがあり、これといった傾向は見つけられなかった。

7. まとめ

今回の天気予測では予測日の前日と当日の気温差や湿度の違いなどの連続性については考慮できなかったが、そういったことも含めて予測を行うと精度が上がるのではないかと考えられる。

