Pembelajaran Mesin

Semester Genap Tahun Akademik 2017-2018

Tugas 1.4: Support Vector Machine

Nadine Azhalia (1301154519)

Support Vector Machine (SVM) merupakan salah satu algoritma *supervised learning*, tujuannya adalah memisahkan kedua kelas menggunakan sebuah garis yang linier. Dalam SVM garis yang dapat memisahkan dua kelas disebut sebagai *hyperlane.* Tujuan dari SVM adalah mencari *hyperlane* yang optimal dengan cara memaksimalkan margin (jarak antar *hypelane* ke data terdekat).

Setiap algoritma memiliki kelebihan dan kekurangan masing-masing. Tergantung pada permasalahan apa yang akan ditangani atau diselesaikan menggunakan algoritma tersebut. Berikut akan dijelaskan kelebihan dan kekurangan dari algoritma *supervised learning*.

1. Jelaskan kelebihan dan kekurangan SVM dibanding Naïve Bayes.

**Jawab:**

Kelebihan:

* Untuk pengklasifikasian dokumen, SVM memiliki tingkat keakuratan yang lebih tinggi dibandingkan dengan Naïve Bayes
* SVM akan bekerja lebih baik daripada NB jika pemilihan parameter tepat.

Kekurangan:

* SVM memiliki perhitungan yang lebih rumit dibandingkan dengan NB. Hal ini karena pada SVM butuh lebih banyak proses *training* untuk hasil yang optimum. Sedangkan pada NB kita hanya membutuhkan nilai *posterior*, *Likelihood*, *Prior*, *Evidence* untuk mencari nilai optimum saat proses *training*.
* Proses *testing* pada NB sangat mudah, hanya melihat tabel atau menghitung nilai probabilitas dengan perkiraan distribusi. Sedangkan pada SVM harus mencari besar margin yang maksimal dengan Lagrangian.

2. Jelaskan kelebihan dan kekurangan SVM dibanding Multi-layer Perceptron.

**Jawab:**

Kelebihan:

* Dari segi prediksi kinerja, SVM secara substansial memiliki kecepatan yang lebih tinggi dibandngkan dengan MLP. Hal ini disebabkan karena pada MLP untuk tiga lapisan (*input layer, hidden layer,* dan *output layer*) memerlukan penggandaan berurutan dari vektor masukan. Sedangkan pada SVM klasifikasi melibatkan penentuan sisi dari batasan keputusan yang diberikan.
* SVM tidak akan terjebak di lokal minimum, SVM akan mencari nilai global minimum. Sedangkan MLP bisa saja terjebak di lokal minimum.
* SVM akan memiliki prediksi akurasi yang lebih baik daripada MLP. Karena SVM mendapatkan pemisah optimum yang memiliki kinerja bagus pada data yang tidak terlihat.

Kekurangan:

* Dari segi waktu *training*, SVM jauh lebih lambat dibandingkan dengan MLP. Hal ini disebabkan karena SVM memerlukan pemecahan masalah Lagrangian.
* MLP memungkinkan untuk memiliki lebih dari satu hasil (*output*) yang disetiap hasilnya memiliki peluang tersendiri. Sedangkan pada SVM hanya memiliki satu hasil saja.
* Pada MLP, data *noise* tidak terlalu diperhatikan. Sedangkan pada SVM data *noise* sangat mempengaruhi untuk akurasi.

3. Jelaskan kelebihan dan kekurangan SVM dibanding Probabilistic Neural Network

**Jawab:**

Kelebihan:

* Pada SVM, parameter yang dibutuhkan hanya parameter untuk pembuatan *hyperlane* dengan persamaan garis liniear. Pada MLP semakin banyak parameter yang digunakan, maka akan semakin rumit jaringan syaraf yang harus dibangun.

Kekurangan:

* Dalam pengklasifikasian, SVM lebih cocok untuk mengklasifikasikan dua kelas. Sedangkan NB bisa dapat mengklasifikasikan lebih dari dua kelas
* Pada PNN kita hanya perlu mengobservasi nilai sigma hingga mendapat akurasi yang tinggi. Sedangkan pada SVM kita harus mencari nilai margin yang maksimal untuk hasil yang akurat.
* Pada PNN, data *training* yang baru dapat ditambahkan atau dihilangkan tanpa harus melakukan *training* ulang. Pada SVM, jika ada data *training* baru yang akan ditambahkan atau dihilangkan maka harus melakukan *training* ulang. Karena jika ada data *training* yang baru pada SVM maka akan mengecohkan *hyperplane*.

Referensi:

[1]Slide Perkuliahan

[2] Ariadi Dio and Kartika Fithriasari. Klasifikasi Berita Indonesia Menggunakan Metode Naïve Bayesian Classification dan Support Vector Machine dengan Confix Stripping Stemmer. JURNAL SAINS DAN SENI ITS Vol. 4, No.2, (2015) 2337-3520.

[3] <https://stats.stackexchange.com/questions/58214/when-does-naive-bayes-perform-better-than-svm>

[4] <http://scikit-learn.org/stable/modules/svm.html>

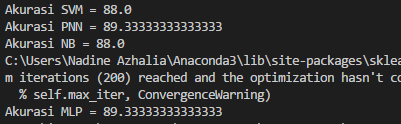
[5] <https://stackoverflow.com/questions/10672268/svm-versus-mlp-neural-network-compared-by-performance-and-prediction-accuracy>

[6] E.A. Zanaty. Support Vector Machines (SVMs) versus Multilayer Perception (MLP) in data classification. Volume 13, Issue 3, November 2012.

[7] <https://stackoverflow.com/questions/11632516/what-are-advantages-of-artificial-neural-networks-over-support-vector-machines>

[8] <http://www.svms.org/anns.html>

Soal Bonus:



Dari hasil pengamatan, dapat diperoleh bahwa algoritma PNN memiliki akurasi yang paling tinggi dibandingkan dengan algoritma SVM, NB, dan MLP untuk kasus pengklasifikasian data yang jumlahnya tidak terlalu banyak.