

PR Pembelajaran Mendalam #1¹

1 Pengantar

PR ini terdiri dari dua buah paket eksperimen. Eksperimen bagian pertama (disingkat E1) relatif mudah, eksperimen bagian kedua (disingkat E2) agak lebih sulit dan memakan waktu lebih lama.

Untuk E1 akan digunakan *Python code* yang sudah dipelajari di kuliah yang lalu untuk memecahkan masalah klasifikasi angka. Di E1 ini akan digunakan 1.797 observasi dengan 64 features (variables) untuk model *Artificial Neural Network (ANN)* dengan dua buah *hidden layers* dengan tiap *hidden layer* terdiri dari 128 neurons (bila perlu neurons boleh ditambah). Di model ini akan ditaksir 26.122 buah parameter. Pada eksperimen E1 ini belum digunakan komponen validasi hasil prediksi. Target yang ingin dicapai di E1 ini berupa hasil training dengan prediksi yang mempunyai akurasi minimal 95%.

Sedangkan untuk E2 akan digunakan *Python code* tentang klasifikasi jenis pakaian yang akan segera dibagikan. Di E2 ini akan digunakan 60.000 observasi untuk training dengan 784 features (variables) untuk model *Artificial Neural Network (ANN)* dengan dua buah hidden layer dengan tiap hidden layer terdiri dari 784 neurons (bila perlu neurons boleh ditambah). Di model ini akan ditaksir lebih dari sejuta parameter, tepatnya 1.238.730 buah parameter. Pada eksperimen E2 ini sudah digunakan komponen validasi hasil prediksi di tahap training berupa penggunaan data test sebanyak 10.000 observasi untuk prediksi. Data untuk test ini berada diluar dari data sebanyak 60.000 observasi yang digunakan untuk training. Langkah validasi dilakukan dengan mengamati akurasi hasil dari penggunaan data untuk training yang akan dibandingkan dengan akurasi hasil penggunaan data untuk test. Target yang ingin dicapai di E2 ini berupa hasil training dengan prediksi yang mempunyai akurasi minimal 85%.

2 Prosedur eksperimen

1. PR ini akan dikerjakan secara tim yang terdiri dari dua mahasiswa. Pada PR ini akan banyak yang akan dikerjakan karena itu sebaiknya masing-masing tim perlu melakukan strategi pembagian tugas untuk mengerjakannya.
2. Untuk E1 dan E2, tiap tim akan menggunakan `seed()` untuk bilangan acak dengan penggabungan kedua NIM mahasiswa. Contohnya, dua buah NIM 10116500 dan 1011600 digabung menjadi `seed(1011650010116600)`. Dengan `seed()` yang

¹Disusun oleh M. Syamsuddin, Prodi Matematika FMIPA-ITB. Versi tanggal 13 Februari 2020.

berbeda-beda ini hasil eksperimen yang diperoleh untuk suatu tim akan berbeda dengan tim yang lain.

3. Untuk E1 dan E2, mula-mula digunakan dulu **epochs** (jumlah iterasi) yang kecil dulu, misalnya 3 atau 5 lalu dimulai dengan *learning rate* **alpha** sebesar 1 lalu dinaikkan atau diturunkan nilainya. Bila sudah diperoleh nilai alpha yang cocok, baru dicoba nilai epochs yang lebih besar.
4. Untuk E1 dan E2, mula-mula digunakan dulu features (variables) tanpa standardisasi (normalisasi). Dua jenis standardisasi (normalisasi) untuk features (variables) seperti yang sudah dibahas di kelas harus dicoba dan dilaporkan hasilnya.
5. Untuk E1 dan E2, berbagai kombinasi fungsi aktivasi (*activation function*) yang harus dicoba dalam setiap eksperimen dan dilaporkan hasilnya berupa:
 - (a) Hidden layers memakai fungsi aktivasi **Sigmoid**, output layer memakai fungsi aktivasi **Softmax** dengan fungsi kerugian (loss function) berupa **cross entropy (log-loss)**. Perhatikan perumusan untuk memperoleh **delta3, delta2, delta1** untuk eksperimen ini.
 - (b) Hidden layers memakai fungsi aktivasi **Tanh**, output layer memakai fungsi aktivasi **Softmax** dengan fungsi kerugian (loss function) berupa **cross entropy (log-loss)**. Perhatikan perumusan untuk memperoleh **delta3, delta2, delta1** untuk eksperimen ini.
 - (c) Hidden layers memakai fungsi aktivasi **Sigmoid**, output layer memakai fungsi aktivasi **Sigmoid** dengan fungsi biaya (cost function) berupa **mean squared errors (MSE)**. Perhatikan perumusan untuk memperoleh **delta3, delta2, delta1** untuk eksperimen ini.
 - (d) Hidden layers memakai fungsi aktivasi **Tanh**, output layer memakai fungsi aktivasi **Sigmoid** dengan fungsi biaya (cost function) berupa **mean squared errors (MSE)**. Perhatikan perumusan untuk memperoleh **delta3, delta2, delta1** untuk eksperimen ini.
6. Tuliskan sebuah laporan untuk tiap tim tentang langkah eksperimen yang ditempuh dan hasil pengamatan eksperimennya. Bila perlu laporan dilengkapi dengan tabel dan gambar. Sertakan pula Python code dalam bentuk Notebook yang bisa dijalankan di JupyterLab. Laporan ini diserahkan lewat japri ke WA saya paling lambat tanggal 25 Februari 2020 jam 12:00.

Selamat bekerja