
Bagian I

1. (Nilai 10) Lingkarilah pernyataan yang benar dari pernyataan berikut ini.

- a. Inferensi pada bagging dilakukan dengan training model dari dataset hasil bootstrap sampling dan dengan agregasi model menggunakan majority voting hasil prediksi semua model.
Salah. Inferensi pada bagging ~~dengan training model dari dataset hasil bootstrap sampling dan~~ dengan agregasi model menggunakan majority voting hasil prediksi semua model.
- b. Inferensi semua base model pada parallel ensemble ataupun sequential ensemble tetap dapat dilakukan secara paralel.
Salah. Inferensi semua base model pada parallel ensemble ataupun sequential ensemble tetap dapat dilakukan secara paralel: bagging ya, voting ya, boosting ya, stacking tidak.
- c. Ensemble diversity pada random forest berupa model dilatih dengan variasi dataset hasil bootstrap sampling dan random subspace.
Benar
- d. Ensemble diversity pada heterogeneous ensembles (with weighting dan with meta learning) berupa model dilatih dengan variasi algoritma learning.
Benar
- e. Ensemble diversity pada Gradient Boosting berupa model dilatih dengan variasi bobot setiap instance pada dataset yang sama.
Salah. Ensemble diversity pada **Gradient Boosting** berupa model dilatih dengan variasi bobot setiap instance pada dataset yang sama. → **seharusnya AdaBoost**
- f. Ensemble diversity pada AdaBoost berupa model dilatih dengan variasi target residual pada setiap iterasi.
Salah. Ensemble diversity pada **AdaBoost** berupa model dilatih dengan variasi target residual pada setiap iterasi. → **seharusnya Gradient Boosting**
- g. Jika diberikan satu ensemble models, agregasi model Adaboost dengan algoritma versi Han et al. (2022) dan algoritma versi Kunapuli (2023) dapat memberikan hasil prediksi yang berbeda.
Benar
- h. Pada AdaBoost versi Kunapuli (2023), instance yang salah diklasifikasikan oleh model pada iterasi t akan memiliki bobot yang lebih besar dari iterasi $t-1$ dengan mengalikan bobot saat ini dengan e^α dengan α menyatakan error pada iterasi t .
Salah. Pada AdaBoost versi Kunapuli (2023), instance yang salah diklasifikasikan oleh model pada iterasi t akan memiliki bobot yang lebih besar dari iterasi $t-1$ dengan

mengalikan bobot saat ini dengan e^{α} dengan α menyatakan **error** pada iterasi t . \rightarrow seharusnya bobot model.

- i. Pada Gradient Boosting dengan y adalah nilai target dan \hat{y} adalah nilai prediksi, $\text{residual} = y - \hat{y}$ merupakan negative gradient dengan menggunakan squared loss function. Nilai residual dan squared loss selalu sama.
Salah. Pada Gradient Boosting dengan y adalah nilai target dan \hat{y} adalah nilai prediksi, $\text{residual} = y - \hat{y}$ merupakan negative gradient dengan menggunakan squared loss function. **Nilai residual dan squared loss selalu sama.**
- j. Hasil prediksi suatu instance pada Adaboost versi Han et al. (2022) dengan n base model dilakukan dengan mengambil hasil prediksi base model terakhir.
Salah. Hasil prediksi suatu instance pada Adaboost versi Han et al. (2022) dengan n base model dilakukan dengan mengambil **hasil prediksi base model terakhir.**

2. (Nilai 12.5) Diberikan model CNN dengan arsitektur sebagai berikut (*complex layer terminology*) untuk klasifikasi gambar:

- input layer berukuran $32 \times 32 \times 3$,
- satu layer konvolusi dengan 5 kernel berukuran 7×7 , dengan padding = 1 dan pergeseran 3 sel, tahap detector dengan ReLU, serta tahap maxpooling dengan kernel berukuran 4×4 , tanpa padding, dan pergeseran 2 sel.
- satu layer *fully connected (dense)* dengan jumlah neuron 10
- output layer 5 neuron dengan fungsi aktivasi softmax

Berikanlah jawaban dengan mencantumkan detail perhitungan.

- a. (Nilai 5) Tentukanlah ukuran feature map yang dihasilkan dari layer konvolusi dengan mengisi tabel berikut ini. Detail perhitungan wajib diisi.

| Tahap | Detail perhitungan | Ukuran Feature Map |
|-------|--------------------|--------------------|
| | | |
| | | |

- b. (Nilai 7.5) Tentukanlah jumlah bobot (trainable parameter) pada model CNN tersebut (shared parameter = True). Output shape diisi dgn dimensi feature map (dalam format $v \times v \times d$) atau dimensi vektor (dalam format length)

| Layer | Detail perhitungan | #Trainable Param | Output Shape |
|-------------|--------------------|------------------|--------------|
| Konvolusi | | | |
| FC (hidden) | | | |

| | | | |
|-----------------------|--|---|--|
| FC (output) | | | |
| Total trainable param | | - | |

3. (Nilai 7.5) Lakukanlah forward propagation untuk mendapatkan feature map dari suatu layer konvolusi jika diberikan data input (\mathbf{x}) berukuran $3 \times 3 \times 2$ pada suatu layer konvolusi yang memiliki 1 kernel berukuran 2×2 . Stride yang digunakan 1, dan tanpa menggunakan padding. Tahap detector dengan ReLU, dan tahap pooling dengan fungsi maksimum dengan kernel berukuran 2×2 , tanpa padding, dan pergeseran 2 sel.
Berikanlah secara eksplisit feature map per tahap dari layer konvolusi tersebut.

Data input \mathbf{x}

| | | |
|---|---|---|
| 1 | 0 | 1 |
| 0 | 0 | 1 |
| 1 | 1 | 0 |

| | | |
|---|---|---|
| 1 | 0 | 1 |
| 0 | 1 | 0 |
| 1 | 0 | 1 |

Kernel (bias 0.5)

| | |
|---|----|
| 1 | 0 |
| 0 | -1 |

| | |
|---|----|
| 0 | -1 |
| 1 | 0 |

Jawab:

1.

- a. **Salah.** Inferensi pada bagging ~~dengan training model dari dataset hasil bootstrap sampling dan~~ dengan agregasi model menggunakan majority voting hasil prediksi semua model.
- b. **Salah.** Inferensi semua base model pada parallel ensemble ataupun sequential ensemble tetap dapat dilakukan secara paralel: bagging ya, voting ya, boosting ya, stacking tidak.
- c. **Benar**
- d. **Benar**
- e. **Salah.** Ensemble diversity pada **Gradient Boosting** berupa model dilatih dengan variasi bobot setiap instance pada dataset yang sama. → **seharusnya AdaBoost**
- f. **Salah.** Ensemble diversity pada **AdaBoost** berupa model dilatih dengan variasi target residual pada setiap iterasi. → **seharusnya Gradient Boosting**
- g. **Benar**
- h. **Salah.** Pada AdaBoost versi Kunapuli (2023), instance yang salah diklasifikasikan oleh model pada iterasi t akan memiliki bobot yang lebih besar dari iterasi $t-1$ dengan mengalikan bobot saat ini dengan e^α dengan α menyatakan **error** pada iterasi t . → **seharusnya bobot model.**
- i. **Salah.** Pada Gradient Boosting dengan y adalah nilai target dan \hat{y} adalah nilai prediksi, $\text{residual} = y - \hat{y}$ merupakan negative gradient dengan menggunakan squared loss function. ~~Nilai residual dan squared loss selalu sama.~~
- j. **Salah.** Hasil prediksi suatu instance pada Adaboost versi Han et al. (2022) dengan n base model dilakukan dengan mengambil **hasil prediksi base model terakhir.**

2a.

| Tahap | Detail perhitungan | Ukuran Feature Map |
|----------------------|--|-----------------------|
| Konvolusi + Detector | $V = 1 + (W - F + 2P) / S = 1 + (32 - 7 + 2 \cdot 1) / 3 = 10$ | $10 \cdot 10 \cdot 5$ |
| Pooling | $V = 1 + (W - F + 2P) / S = 1 + (10 - 4 + 0) / 2 = 4$ | $4 \cdot 4 \cdot 5$ |

2b.

| Layer | Detail perhitungan | #Trainable Param | Output Shape |
|-----------------------|------------------------------------|------------------|---------------------|
| Konvolusi | $5 \cdot (7 \cdot 7 \cdot 3 + 1)$ | 740 | $4 \cdot 4 \cdot 5$ |
| FC (hidden) | $(4 \cdot 4 \cdot 5 + 1) \cdot 10$ | 810 | 10 |
| FC (output) | $(10 + 1) \cdot 5$ | 55 | 5 |
| Total trainable param | | 1605 | - |

3.

Jawab:

Tahap konvolusi

| | |
|---------------------------------|---------------------------------|
| $1+0+0+0+0+0+0+0=1+0.5=1.5$ | $0+0+0+-1+0+-1+1+0=-1+0.5=-0.5$ |
| $0+0+0+-1+0+-1+1+0=-1+0.5=-0.5$ | $0+0+0+0+0+0+0+0=0+0.5=0.5$ |

Tahap detector

| | |
|-----|-----|
| 1.5 | 0 |
| 0 | 0.5 |

Tahap maxpooling

| |
|-----|
| 1.5 |
|-----|

Tuliskan jawaban di bagian kosong di halaman ini atau di baliknya.

Bagian II

Untuk soal 1 dan 2: Diberikan 2 data latih dengan 2 fitur x_1 dan x_2 , serta label kelas y di bagian kanan. Tentukan jenis pembelajaran Perceptron yang tepat untuk setiap kasus data tersebut. Jawaban mungkin lebih dari satu. (Nilai: 4)

- (a) Perceptron training rule
(c) Stochastic gradient descent

- (b) Batch gradient descent
(d) Tidak ada yang benar

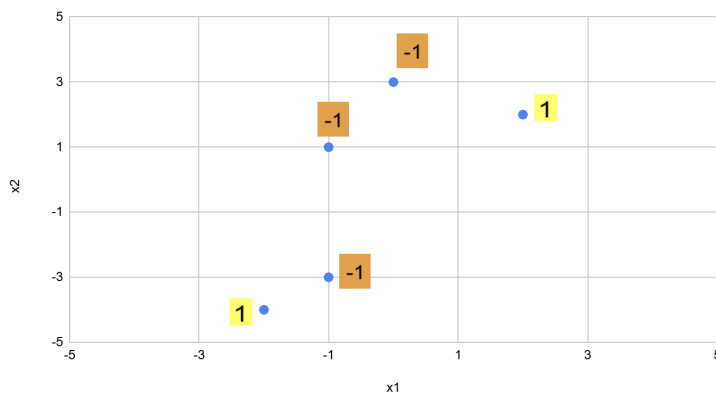
1. Jawaban: a, b, c

Alasan pilihan:

Data bersifat linearly separable, ketiga jenis algoritma dapat digunakan.

| | | | | |
|-----------|---|----|----|----|
| x1 | 1 | -1 | -2 | -1 |
| x2 | 2 | -1 | 1 | -3 |
| y | 1 | -1 | 1 | -1 |

Soal 2



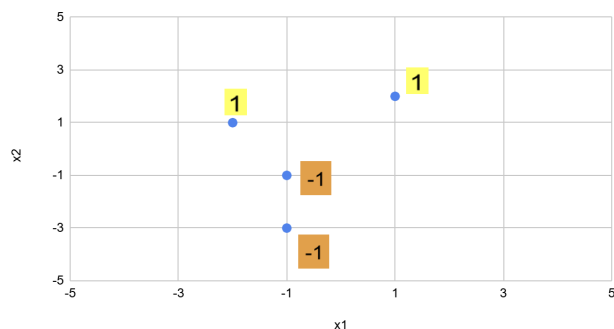
2. Jawaban: b, c

Alasan pilihan:

Data bersifat not linearly separable, hanya algoritma batch dan stochastic gradient descent yang dapat digunakan.

| | | | | | |
|-----------|---|----|----|----|----|
| x1 | 2 | 0 | -1 | -1 | -2 |
| x2 | 2 | 3 | 1 | -3 | -4 |
| y | 1 | -1 | -1 | -1 | 1 |

Soal 1



Untuk soal 3 s.d. 6: Tentukan apakah pernyataan berikut Benar atau Salah, dan tuliskan alasan jawaban tsb. (Nilai: 6)

3. Pada pelatihan dengan menggunakan Perceptron Training Rule, update bobot dilakukan setiap sepanjang iterasi.

Jawaban:

BENAR, sesuai algoritma PTR.

4. Prinsip algoritma pembelajaran Gradient Descent pada Perceptron adalah merupakan strategi pencarian dengan kemiringan untuk mendapatkan error nol.

Jawaban:

SALAH, seharusnya untuk mendapatkan error minimum.

5. Pada Perceptron Training Rule dan Stochastic Gradient Descent proses update bobot dilakukan dengan mekanisme dan aturan yang sama.

Jawaban:

SALAH, seharusnya berbeda, yaitu pada PTR update hanya dilakukan jika $o_i \neq y_i$.

6. Penggunaan learning rate yang yang besar pada Perceptron pasti konvergen ke lokal minimum.

Jawaban:

SALAH, learning rate yang kecil lebih cenderung ke lokal minimum, tapi lebih lambat

7. Terdapat dataset yang terdiri atas 2 atribut berikut ini, beserta label kelas y. Encoding kelas positif adalah 1 dan kelas negatif adalah -1, dengan batas untuk kelas positif > 0 . Inisialisasi bobot semua atribut adalah 1, dan inisialisasi bobot bias adalah -1. Learning rate 0.1. Akan dilakukan pelatihan model Perceptron sebanyak 1 epoch. (Nilai: 15)

| Data Latih | | | | | Data Uji | | | |
|------------|----|----|----|--|----------|-----|------|----|
| id | x1 | x2 | y | | id | x1 | x2 | y |
| d1 | 1 | 0 | 1 | | d4 | 0 | 0.95 | 1 |
| d2 | 0 | 0 | -1 | | d5 | 0.5 | 0 | -1 |
| d3 | 1 | 1 | 1 | | | | | |

- 1) Lakukan proses pembelajaran model dengan **Batch Gradient Descent**. Tuliskan bobot yang dihasilkan dan error dari hasil pembelajaran. [Nilai 5]
- 2) Prediksi kelas untuk data uji yang disediakan menggunakan model yang dihasilkan pada bagian a.1). Hitung akurasi model. [Nilai 2]

| Batch Gradient Descent | | | | | | | | | | | | | |
|------------------------|-----------------|-----------------|-----------------|----------------|----------------|----------------|----------|---------|-----------------|------------------|------------------|------------------|-------|
| | w _{i0} | w _{i1} | w _{i2} | x ₀ | f ₁ | f ₂ | Target | Output | Target – Output | dw _{i0} | dw _{i1} | dw _{i2} | Error |
| | -1 | 1 | 1 | | | | | | | 0 | 0 | 0 | 0 |
| Epoch 1 | -1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0.1 | 0.1 | 0 | 0.5 |
| | -1 | 1 | 1 | 1 | 0 | 0 | -1 | -1 | 0 | 0.1 | 0.1 | 0 | 0.5 |
| | -1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 0 | 0.1 | 0.1 | 0 | 0.5 |
| | | | | | | | | | Total | 0.1 | 0.1 | 0 | 0.5 |
| Update bobot | -0.9 | 1.1 | 1 | | | | | | | 0 | 0 | 0 | 0 |
| Bobot akhir | -0.9 | 1.1 | 1 | | | | | | | | | | |
| | Data baru | | | | f ₁ | f ₂ | sentimen | output | Kelas | | | | |
| | -0.9 | 1.1 | 1 | 1 | 0 | 0.95 | 1 | 0.05 | 1 | | | | |
| | -0.9 | 1.1 | 1 | 1 | 0.5 | 0 | -1 | -0.35 | -1 | | | | |
| | | | | | | | | Akurasi | 100% | | | | |

- b. 1) Lakukan proses pembelajaran model dengan **Stochastic Gradient Descent**. Tuliskan bobot yang dihasilkan dan error dari hasil pembelajaran. [Nilai 5]
2) Prediksi kelas untuk data uji yang disediakan menggunakan model yang dihasilkan pada bagian b.1). Hitung akurasi model. [Nilai 2]

| Stochastic Gradient Descent | | | | | | | | | | | | | | |
|-----------------------------|-----------------|-----------------|-----------------|----------------|----------------|----------------|----------|---------|-----------------|------------------|------------------|------------------|---------|---|
| | w _{i0} | w _{i1} | w _{i2} | x ₀ | f ₁ | f ₂ | Target | Output | Target – Output | dw _{i0} | dw _{i1} | dw _{i2} | Error | |
| | -1 | 1 | 1 | | | | | | | | | | | 0 |
| Epoch 1 | -1 | 1 | 1 | 1 | 1 | 0 | 1 | 0 | 1 | 0.1 | 0.1 | 0 | 0.5 | |
| | -0.9 | 1.1 | 1 | 1 | 0 | 0 | -1 | -0.9 | -0.1 | -0.01 | 0 | 0 | 0.505 | |
| | -0.91 | 1.1 | 1 | 1 | 1 | 1 | 1 | 1.19 | -0.19 | -0.019 | -0.019 | -0.019 | 0.52305 | |
| | | | | | | | | | Total | | | | 0.52305 | |
| Update bobot | -0.929 | 1.081 | 0.981 | | | | | | | 0 | 0 | 0 | 0 | |
| Bobot akhir | -0.929 | 1.081 | 0.981 | | | | | | | | | | | |
| | Data baru | | | | f ₁ | f ₂ | sentimen | output | Kelas | | | | | |
| | -0.929 | 1.081 | 0.981 | 1 | 0 | 0.95 | 1 | 0.00295 | 1 | | | | | |
| | -0.929 | 1.081 | 0.981 | 1 | 0.5 | 0 | -1 | -0.3885 | -1 | | | | | |
| | | | | | | | | Akurasi | 100% | | | | | |

- c. Model mana yang menghasilkan kinerja yang lebih baik, tuliskan alasannya. [Nilai 1]

Kedua model BGD dan SGD memiliki kinerja sama baiknya yaitu 100%.

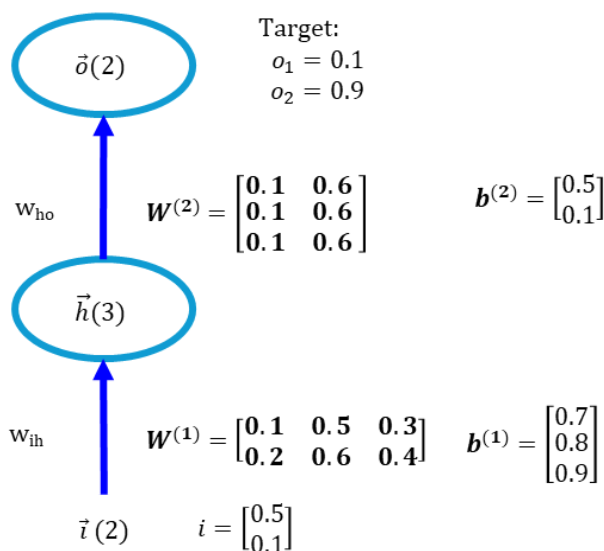
8. Tuliskan minimal 5 perbedaan prinsip Convolutional Neural Network (CNN) dengan FFNN. Berikan penjelasan singkat untuk setiap poin, boleh menambahkan ilustrasi. (Nilai 10)

| Poin Perbedaan | FFNN | CNN |
|-------------------------------|---|---|
| 1. Ekstraksi fitur | Tidak ada ekstraksi fitur secara khusus | Ekstraksi fitur otomatis melalui filter |
| 2. Fitur spasial | Tidak menangani fitur spasial | Dapat menangkap fitur spasial |
| 3. Penanganan data gambar | Tidak dirancang scalable untuk data image besar | Memiliki kemampuan generalisasi dan lebih cepat untuk data gambar |
| 4. Keterhubungan antar neuron | Full connectivity | Local connectivity |
| 5. Parameter sharing | Tidak ada parameter sharing | Memungkinkan parameter sharing |

Bagian III

Terdapat struktur Feed Forward Neural Network (Fully Connected) seperti pada gambar berikut. Fungsi aktivasi untuk *neuron* di *hidden layer* adalah fungsi linear. Fungsi aktivasi untuk *neuron* di *output layer* adalah sigmoid. Fungsi *error* atau *loss*: $\frac{1}{2} * (\text{target} - \text{output})^2$; dengan turunannya terhadap output adalah $-(\text{target} - \text{output})$. Turunan fungsi linear adalah 1. Turunan fungsi sigmoid: $\text{output} * (1 - \text{output})$. Learning rate adalah 0,5. Penamaan bobot: $w_{[\text{asal}][\text{tujuan}]}$, contoh w_{i1h2} artinya bobot dari input 1 ke neuron 2 di *hidden layer*.

Bagian ini terdiri atas 5 soal, jawablah pada tempat yang disediakan. Jawaban yang dituliskan di lembar Bagian II atau Bagian III lain tidak akan dinilai.



| Nomor | Pertanyaan | Jawaban |
|-------|---|---------|
| 1 | Tentukan berapa banyak parameter yang perlu dipelajari berdasarkan arsitektur FFNN pada gambar. Tuliskan angka yang terlibat dan tidak hanya hasil akhir saja. (Nilai 6) | |
| 2 | Tentukan nilai semua neuron pada hidden layer , dan nilai pada output o2 . Jawaban tidak hanya hasil akhir berupa angka, tuliskan semua angka yang terlibat dalam perhitungan. (Nilai 6) | |

| Nomor | Pertanyaan | Jawaban |
|-------|--|---------|
| 3 | Tentukan nilai $\frac{\partial E}{\partial_{net-o1}}$ (turunan error terhadap kombinasi linear o1) dan nilai $\frac{\partial E}{\partial_{net-o2}}$ dengan menuliskan lengkap formula serta semua angka yang terlibat. (Nilai 6) | |
| 4 | Tentukan nilai $\frac{\partial E}{\partial_{net-h3}}$ (turunan error terhadap kombinasi linear h1) dengan menuliskan lengkap formula serta semua angka yang terlibat. (Nilai 6) | |
| 5 | Tentukan nilai bobot w_{i1h3} yang baru dengan menuliskan lengkap formula serta angka yang terlibat. (Nilai 6) | |