

# Ujian Tengah Semester

## TF4063 Sains Data dan Rekayasa

14 Oktober 2020

1. Jelaskan mengenai tiga cara pendekatan berikut untuk menentukan parameter model pada mesin pembelajar:
  - *loss function minimization*
  - *maximum likelihood*
  - *Bayesian approach*
2. (Kode Python) Buatlah data sintetik dengan menggunakan polinom:  $f(x) = 5x^3 - x^2 + x$  dan diberikan noise Gaussian dengan  $\mu = 0$  dan  $\sigma^2 = 300$ . Ambil 100 nilai  $x$  dari distribusi acak seragam dalam rentang  $-5 < x < 5$ . Gunakan hanya data dengan  $x < 0$  dan  $x > 2$ , artinya data yang berada pada  $0 \geq x \geq 2$  akan dibuang. Gunakan pendekatan *maximum likelihood* untuk melakukan pemodelan pada data ini dengan menggunakan polinom dengan orde 1 sampai dengan 8. Plot hasil yang Anda peroleh beserta error bar yang terkait. Lengkapi jawaban Anda dengan kode Python dan persamaan terkait yang digunakan (tidak perlu penurunan). Hasil yang Anda dapatkan akan mereproduksi Gambar 2.17 pada buku Girolami-Rogers.
3. Jelaskan mengenai empat distribusi probabilitas yang terlibat pada aturan Bayes (dalam kaitannya dengan penentuan parameter model)

$$p(\mathbf{w}|\mathbf{t}) = \frac{p(\mathbf{t}|\mathbf{w})p(\mathbf{w})}{p(\mathbf{t})} \quad (1)$$

4. Jelaskan secara singkat tiga cara (aproksimasi) yang dapat dilakukan untuk menentukan parameter model dengan pendekatan Bayesian jika distribusi yang terlibat bukan konjugat.