ECE 445 (ML for ENGG): Homework #1

Waheed U. Bajwa (waheed.bajwa@rutgers.edu)

1. Give an example each of a system that is:
   - A machine learning system.
   - An artificial intelligence system that is *not* a machine learning system.
   - A data science system that is neither a machine learning system nor an artificial intelligence system.
2. What is the difference between the cumulative distribution function (CDF) and the probability density function (PDF) of a scalar-valued random variable (assuming the PDF exists)?
3. Give an example each of a:
   - Linear function
   - Nonlinear function
   - Polynomial function
4. Consider a random variable X that has probability density function $p_X(x)$. Define the *mean*, *median*, and *mode* of this random variable in terms of the PDF $p_X(x)$.
5. Consider a random variable X that is normally distributed (i.e., Gaussian distribution) with mean 1 and variance 0.25. Provide:
   - Mathematical expression for the PDF of X.
   - A rough sketch of the PDF of X.
   - Highlight the mean, median, and mode of X on its PDF.
6. Consider a function $e = f(\theta)$, which is given by $f(\theta) = a\theta^2 + b\theta + c$ for some real numbers a, b, and c. Solve for $\theta^* = \arg\min_\theta(e)$; that is, the value of θ that minimizes e. *Note:* This is a toy example of optimization, which lies at the heart of machine learning algorithms; unlike this problem, real machine learning problems often do not have closed-form solutions to the optimization problems.
7. Consider a random variable Y that is defined as $Y = a + W$, where a is a constant and W is a random variable with mean 0 and variance 1. Suppose we are given N independent realizations of Y, given by $Y_i, i = 1, \dots, N$. Show that $\widehat{a} = \frac{1}{N}\sum_{i=1}^{N} Y_i$ has mean a and variance that satisfies $\lim_{N\to\infty} \mathrm{var}(\widehat{a}) = 0$. *Note:* This is an example of the *Law of Large Numbers (LLN)*, which essentially says that averaging a large number of independent random variables gives us an almost deterministic quantity.
8. What are the conditions under which a matrix admits eigenvalue decomposition?
9. Consider a matrix $\mathbf{A} \in \mathbb{R}^{m\times n}$. What is the relationship between the singular value decomposition of $\mathbf{A}$ and the eigenvalue decomposition of $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$.
10. Provide a mathematical description of the concept of *linearly independent vectors*; that is, when are a given number of vectors (say M) in $\mathbb{R}^m$ linearly independent?