# Assignment 1
# Machine Learning for Business Analytics

Due: October/ 05 23:59 (GMT+8)

**Instruction**

- There are two parts in this assignment. You only have to hand in Part 1 regarding the Julia Programming. **Part 2 will not be graded**; of notice is that similar questions may show up in midterm or final.
- Please hand in your Julia code with a **PDF file**, following the sample code in the **HW1.ipynb** file. Wrong format will not be graded.
- **Copying the assignment will result in zero points.**
- Late submission will be graded according to the following rule:
  your grade = original grade × (1 − 0.05$h$), if you submit $h$ hours after the deadline.

## Part 1 Programming Exercises (100%)

1. (Inner Product) Define a function called "Inner_Product" that produces the result of the inner product of two vectors.
   Example: a = [1, 2, 3], b = [2, 3, 4], your function needs to output 20.
   Input: a, b
   Output: 20

2. (Exception Handling) In the lecture, we have learnt that you cannot apply inner product to two vectors with different dimensions. Define a function called "Strict_inner_Product": if two vectors have the same dimension, output the result of their inner product; otherwise, output a warning and the dimension of two vectors.
   Example:
   (1) a = [1, 2, 3], b = [2, 3, 4, 5], your function should output "Warning! 3*1 vector can't do inner product with a 4*1 vector!"
       Input: a, b
       Output: Warning! 3*1 vector can't do inner product with a 4*1 vector!
   (2) a = [1, 2, 3], b = [2, 3, 4], your function should print 20.
       Input: a, b
       Output: 20

3. (Advanced Exception Handling) In the tutorial, TA taught how to identify the datatype of a variable. If there is a naughty student, named NS, who changes your variable, from an assigned vector to a string or a tuple, the previous functions won't work under NS's trick. Please save TA from NA's trick by creating a new function called "Identify_Wrong_Datatype" which will identity the datatype of your input, if they are allowed

to do inner product, then you output the result of the operation; otherwise, you output the datatype of input.
Example:
(1) a = [1, 2, 3], b = [2, 3, 4], your function should print 20.
Input: a, b
Output: 20
(2) a = “[1, 2, 3]”, b = [1, 2, 3], your function should output “Warning! String(datatype of a) can’t do inner product with Vector(datatype of b)!”
Input: a, b
Output: Warning! String can’t do inner product with Vector!

4. (Eugene’s calculator) Eugene has a magic calculator. When you key in an integer (n) and the operands, it will randomly generate “n” numbers and output the result of mathematical operation with the input operands. The teaching team (TT) thinks it’s so cool to have that calculator. As a result, Eugene wants to give TT the magic calculator as a gift. Yet, TT can’t reproduce the software inside the magic calculator.
The logic of the calculator is as following:
    a. Randomly generate n, the key-in integer, numbers from Normal distribution with mean equals to 0 and variance equals to 100.
    b. Use the operands in the given vector in order to do the mathematical operation.
    c. Output the result
Please help Eugene to develop the software and name it “Gift”.
Example:
(1) Input: n = 4, operands = [‘+’, ‘-’, ‘*’]
Output: 10 (suppose the generated number is a,b,c,d, and the result of (((a+b)-c)*d) is 10)
*Please don’t set random seed inside your function!

5. (Sunny’s Crazy Idea) When Sunny claims reimbursement, sometimes he finds there are some missing values on his sheet. With those missing values, he can’t get the total expense he should receive easily by doing inner product. To avoid errors, he asks his administrative assistant (AA) to create a new function called “Account_Manager”, which will automatically fill the missing value by replicating the value(s) from the top of the vector. Moreover, to be careful, you also need to tell Sunny which value(s) is/are missing and what value do you use to substitute the missing value. The information he has are the names of equipment he bought, their quantities and their prices.
Example:
(1) name = [‘Pencil’, ‘Marker’, ‘Glue’], quantity = [3, 4], price = [30, 50, 80]. The total expense will be 3*30 + 4*50 + 3*30 = 530. Then you should print “The total expense is 530. The quantity for Glue is missing and filled with 3.”
Input: name, quantity, price
Output: The total expense is 530.

The quantity for Glue is missing and filled with 3.

(2) name = [‘Pencil’, ‘Marker’, ‘Glue’’], quantity = [3, 4, 5], price = [30, 50]. The total expense will be 3*30 + 4*50 + 5*30 = 440. Then you should print “The total expense is 440. The price for Glue is missing and filled with 30.”
Input: name, quantity, price
Output: The total expense is 440.
The price for Glue is missing and filled with 30.

(3) name = [‘Pencil’, ‘Marker’, ‘Glue’, ‘Scissor’], quantity = [3, 4], price = [30, 50, 80]. 3*30 + 4*50 + 3*80 + 4*30 = 650. Then you should print “The total expense is 650. The quantity for Glue is missing and filled with 3. The quantity for Scissor is missing and filled with 4. The price for Scissor is missing and filled with 30.”
Input: name, quantity, price
Output: The total expense is 650.
The quantity for Glue is missing and filled with 3.
The quantity for Scissor is missing and filled with 4.
The price for Scissor is missing and filled with 30.

**Part 2 Mathematical Exercises (0%)**
**Chapter 1 Exercises**

1. How many bytes does it take to store 100 vectors of length 105?
How many flops does it take to form a linear combination of them (with 100 nonzero coefficients)?
How long would this take on a computer capable of carrying out 1 Gflop/s?

2. Linear combinations of linear combinations. Suppose that each of the vectors $b_1,\dots, b_k$ is a linear combination of the vectors $a_1,\dots, a_m$, and $c$ is a linear combination of $b_1,\dots, b_k$. Then $c$ is a linear combination of $a_1,\dots, a_m$. Show this for the case with $m = k = 2$.

**Chapter 2 Exercises**

1. *Taylor approximation.* Consider the function $f: \mathbf{R}^2 \to \mathbf{R}$ given by $f(x_1, x_2) = x_1x_2$. Find the Taylor approximation $\hat{f}$ at the point $z = (1, 1)$. Compare $\hat{f}(x)$ and $\hat{f}(x)$ for the following values of $x$:
$x = (1, 1)$
$x = (1.05, 0.95)$
$x = (0.85, 1.25)$
$x = (-1, 2)$.

2. *Affine function*. Suppose $\psi : \mathbf{R}^2 \rightarrow \mathbf{R}$ is an affine function, with $\psi(1, 0) = 1$ and $\psi(1, -2) = 2$. (a) What can you say about $\psi(1, -1)$? Either give the value of $\psi(1, -1)$, or state that it cannot be determined.

   (b) What can you say about $\psi(2, -2)$? Either give the value of $\psi(2, -2)$, or state that it cannot be determined. Justify your answers.

3. Integral and derivative of polynomial. Suppose the *n*-vector *c* gives the coefficients of a polynomial $p(x) = c_1 + c_2 x + \cdots + c_n x^{n-1}$

   (a) Let $\alpha$ and $\beta$ be numbers with $\alpha < \beta$. Find an *n*-vector a for which

   $$a^T c = \int_{\alpha}^{\beta} p(x)\, dx$$

   always holds. This means that the integral of a polynomial over an interval is a linear function of its coefficients.

   (b) Let $\alpha$ be a number. Find an *n*-vector *b* for which

   $$b^T c = p'(\alpha).$$

   This means that the derivative of the polynomial at a given point is a linear function of its coefficients.