

# DSCI 417 – Homework 05

Sean Graham

## Problem 1: Sparse Vector Representation

- 1. [8, [0,3,4], [1,1,1]]
- 2. [10, [2,4,8], [4,7,2]]
- 3. [10, [2,8], [1,1]]
- 4. [3, [0,1,2], [4,1,8]]
- 5. [6, [], []]

- 1. [0, 0, 1, 0, 0, 3, 0, 9]
- 2. [0, 0, 0, 0]
- 3. [0, 2, 3, 1]
- 4. [0, 1, 0, 0, 0, 1, 0, 0]
- 5. [0, 0, 0, 0, 2, 0]

## Problem 2: One-Hot Encoding

Group	Group_ix	Group_vec
C	1	[8, [1], [1]]
E	2	[8, [2], [1]]
G	4	[8, [4], [1]]
F	0	[8, [0], [1]]
A	6	[8, [6], [1]]
D	7	[8, [7], [1]]

Problem 3: Vector Assembler (Part 1)

c1	c2	c3	features
3	B	Y	[8,[0,1,7],[3,1,1]]
6	C	X	[8,[0,4,6],[6,1,1]]
2	D	X	[8,[0,3,6],[2,1,1]]
5	A	Y	[8,[0,2,7],[5,1,1]]
1	A	Z	[8,[0,2,5],[1,1,1]]

Problem 4: Vector Assembler (Part 2)

c1	c2	c3	features
1	A	X	(8,[0,2,6],[1,1,1])
4	D	Y	(8,[0,3,7],[4,1,1])
5	B	X	(8,[0,1,6],[5,1,1])
7	C	Z	(8,[0,4,5],[7,1,1])
1	C	Y	(8,[0,4,7],[1,1,1])