1.1.

$$\frac{\partial NLL}{\partial w}=-\sum_{i=1}^{N}[-\frac{1-y_i}{1-\sigma(w^Tx_i)}+\frac{y_i}{\sigma(w^Tx_i)}]\frac{\partial\sigma(w^Tx_i)}{\partial w}$$

$$=-\sum_{i=1}^{N}[-\frac{1-y_i}{1-\sigma(w^Tx_i)}+\frac{y_i}{\sigma(w^Tx_i)}](1-\sigma(w^Tx_i))\sigma(w^Tx_i)x_i$$

$$=\frac{-\sum_{i=1}^{N}[\sigma(w^Tx_i)(1-y_i)+y_i-y_i\sigma(w^Tx_i)}{(1-\sigma(w^Tx_i))\sigma(w^Tx_i)}](1-\sigma(w^Tx_i))\sigma(w^Tx_i)x_i$$

$$=-\sum_{i=1}^{N}x_i(y_i-\sigma(w^Tx_i))$$

1.2

a.

$$l(w)=(1-y_t)\log(1-\sigma(w^Tx_t))+y_t\log\sigma(w^Tx_t)$$

b.

$$w_t=w_{t-1}+\eta x_t(y_t-\sigma(w^Tx_t))$$

c. The time complexity comes from dot product, It take O(d) for the dot product and there are n such products, but assuming x_t being sparse. $d\simeq 1$, then, it would be o(n)

d. the given equation is equivalent to:

$$\frac{\partial l}{\partial w}-\mu\frac{\partial\|w\|^2}{\partial w}=\frac{\partial l}{\partial w}-2\mu w=-\sum_{i=1}^{N}x_i(y_i-\sigma(w^Tx_i))-2\mu w$$

$$w_t=w_{t-1}+\eta\left(x_t(-\sigma(w_t^Tx_t)+y_t)-2\eta\mu w_{t-1}\right)$$

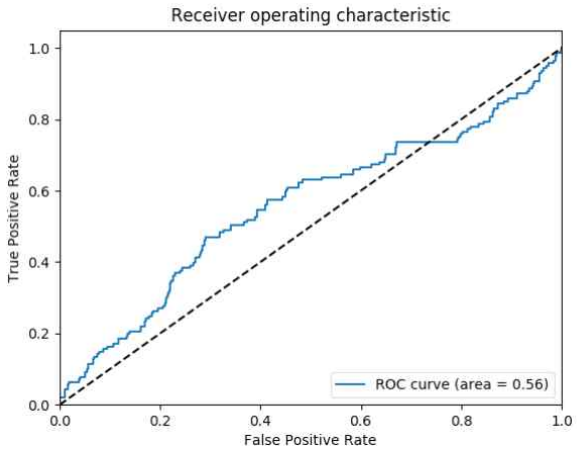
For each update, the time complexity will be O(d)

2.1

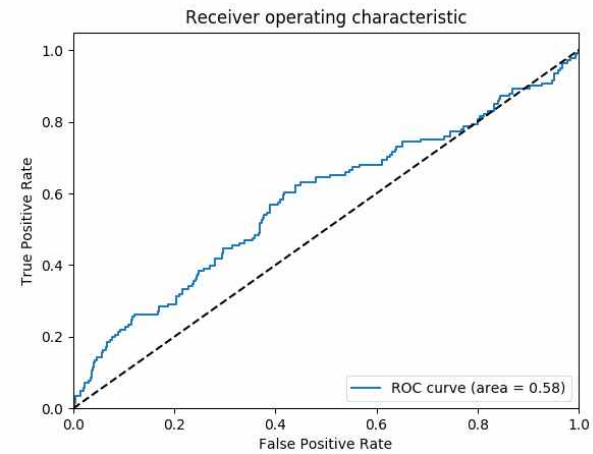
	Deaseased		Alive	
Event Count				
1.Avg Event Count	1027.738		683.155	
2.Max Event Count	16829		12627	
3.Min Event Count	2		1	
Encounter Count				
1.Avg Encounter Count	24.839		18.695	
2.Median Encounter Count	14		9.0	
2.Max Encounter Count	375		391	
3.Min Encounter Count	1		1	
Record Length				
1.Avg Record Length	157.042		194.703	
2.Median Record Length	25.0		16.0	
3.Max Record Length	5364		3103	
4.Min Record Length	0		0	
Common Diagnosis	DIAG320128	416	DIAG320128	1018
	DIAG319835	413	DIAG319835	721
	DIAG313217	377	DIAG317576	719
	DIAG197320	346	DIAG42872402	674
	DIAG132797	297	DIAG313217	641
Common Lab Test	LAB3009542	32765	LAB3009542	66937
	LAB3023103	28395	LAB3000963	57751
	LAB3000963	28308	LAB3023103	57022
	LAB3018572	27383	LAB3018572	54721
	LAB3016723	27060	LAB3007461	53560
Common Medication	DRUG19095164	6396	DRUG19095164	12468
	DRUG43012825	5451	DRUG43012825	10389
	DRUG19049105	4326	DRUG19049105	9351
	DRUG956874	3962	DRUG19122121	7586
	DRUG19122121	3910	DRUG956874	7301

2.2.b

$\mu = 0.0, \eta = 0.01$



$\mu = 0.1, \eta = 0.01$



$\mu = 0.1, \eta = 0.06$