Problem set 5
Usvyatsov Mikhail

Problem 1

$MI(X,Y) = H(Y) - H(Y|X) = -\sum_y p(y)ln(p(y)) + \sum_x p(x) \sum_y p(y|x)ln(p(y|x)) =$
$= -\sum_y p(y)ln(p(y)) + \sum_{x,y} p(y|x)p(x)ln(p(y|x)) = \sum_{x,y} p(x,y)ln(\frac{p(x,y)}{p(x)}) - \sum_y p(y)ln(p(y)) =$
$= \sum_{x,y} p(x,y)ln(\frac{p(x,y)}{p(x)}) - \sum_{x,y} p(y|x)ln(p(y)) = \sum_{x,y} p(x,y)ln\left(\frac{p(x,y)}{p(x)p(y)}\right)$
The last formula is exactly $KL(p||g)$. QED.

Problem 2

$K(x,z) = e^{-\gamma||x-z||^2}$
According to Merser's theorem function k is a kernel iff:

1. K(x, z) = K(z, x), that obviously holds in our case.
2. $\int_x \int_z K(x,z)g(x)g(z)dxdz \geq 0$ for every function g: X $\rightarrow$ R

To prove the second condition is quite difficult.
Instead of this we can see that:
$e^{-\gamma||x-z||^2} = \sum_{k=0}^{\infty} \frac{(2\gamma<x,z>)^k}{k!}$
That is the sum of products of kernels $2\gamma < x, z >$. According to the materials of lecture 9 $< x, z >$ is a valid Mercer Kernel.

Due to the fact that given two Mercer kernels its multiplication is also a Mercer Kernel and its sum is also a valid Mercer Kernel. QED.

Problem 3

$\omega_i^{m+1} = \frac{\omega_i^m e^{-\alpha_m y_i h^m(x_i)}}{Z_m}$
$\sum_i \omega_i^{m+1} = 1$
From previous we can derive: $Z_m = \sum_i \omega_i^{m+1} = \sum_i \omega_i^m e^{-\alpha_m y_i h^m(x_i)} = \sum_{correct} \omega_i^m e^{-\alpha_m y_i h^m(x_i)} + \sum_{incorrect} \omega_i^m e^{-\alpha_m y_i h^m(x_i)} = \epsilon_m e^{\alpha_m} + (1-\epsilon_m)e^{-\alpha_m}$, where $\alpha_m = \frac{1}{2}log\left(\frac{1-\epsilon_m}{\epsilon_m}\right)$

Then updated error rate is $\epsilon_{m+1} = \sum_{incorrect} \omega_i^{m+1} = \sum_{incorrect} \frac{\omega_i^m e^{-\alpha_m y_i h^m(x_i)}}{Z_i} = \sum_{incorrect} \frac{\omega_i^m e^{-\alpha_m}}{Z_i} = \frac{1}{2}$