EL7006 Redes Neuronales y Teoría de Información para el Aprendizaje
Tarea 2

Profesor: Pablo Estévez
Auxiliar: Pablo Huijse
Ayudante: Pablo Huentelemu
Semestre: Primavera 2014

Considere un modelo

$$Z(x) = ax + bx^2 + \nu,$$

donde ν es la componente asociada al ruido y $Z(x)$ son las observaciones. Usted tiene a su disposición los pares $\{(x_i, z_i)\}_{i=1,\dots,N}$. El objetivo es entrenar un aproximador parámetrico $g(x; w) = w_1 x + w_2 x^2$ para recuperar los parámetros del sistema (a y b). La función de densidad de probabilidad (fdp) del ruido está dada por $p_\nu(\nu) = 0.8\,\mathcal{N}(0, 0.1) + 0.2\,\mathcal{N}(10, 0.1)$, que corresponde a una mezcla impulsiva de Gaussianas (modelo Middleton). Entrene un regresor usando como función de costo el MEE (Minimización de Entropía del Error). Esto es equivalente a maximizar el potencial de información

$$\text{máx}\, V = \frac{1}{N^2} \sum_{i=1}^{N} \sum_{j=1}^{N} G_\sigma \left(e_i - e_j\right),$$

donde $G_\sigma(\cdot)$ es un kernel Gaussiano con ancho de banda σ.

Atención: Para resolver esta tarea no es necesario programar. Utilice el *ipython notebook*[1] que viene adjunto a este documento, el cual tiene en cada uno de sus bloques los contenidos de la tarea previamente programados. Al final de este documento encontrará instrucciones generales para usar el *ipython notebook*.

A continuación se detalla paso a paso lo que usted debe realizar.

1. Genere las observaciones $Z(x)$. Considere como parámetros del sistema $b = -2.0$ y $a = D$, donde D es el dígito verificador de RUT (Para $D = K$ use $a = 10$).

2. Entrene un aproximador paramétrico usando gradiente descendente y el MSE como función de costo. Considere un aproximador con dos pesos definido como $g(x; w) = w_1 x + w_2 x^2$. Utilice 1000 épocas y una tasa de aprendizaje $\mu = 0.0002$. Estudie la evolución del error intrínseco $\mathbb{E}\left[(g(x; w) - f(x))^2\right]$ y visualice los resultados del regresor en el espacio de entrada. Describa lo que observa justificando apropiadamente. Compare los pesos obtenidos por el regresor con los parámetros intrínsecos del sistema. Considere estos resultados como referencia para comparar con el regresor MEE.

3. Entrene un aproximador paramétrico usando gradiente descendente y el MEE como función de costo. Considere un aproximador equivalente al usado en el punto anterior. Entrene el aproximador usando MSE durante 200 épocas. Luego entrene usando MEE por 800 épocas a partir de los resultados del anterior. Utilice una tasa de aprendizaje $\mu = 1.0$. Pruebe los siguientes valores para el ancho de banda del kernel Gaussiano: $\sigma = 0.2,\ 2.0,\ y\ 20.0$. Describa lo que observa justificando apropiadamente. Compare los pesos obtenidos por los regresores MEE con los parámetros intrínsecos del sistema. ¿Con cuál valor de σ se obtiene un mejor desempeño (mínimo error intrínseco

[1] `http://ipython.org/notebook.html`

al final del entrenamiento)? Cuál es la función del ancho de banda? ¿Qué criterios existen para seleccionar este parámetro? Justifique sus respuesta basándose en la teoría.

4. Compare el mejor regresor MEE con el regresor MSE del punto (2) en términos de la evolución del error intrínseco, parámetros obtenidos y calidad del regresor. ¿Cuál regresor es más robusto al ruido impulsivo? Explique porqué el criterio MEE obtiene un mejor desempeño que el criterio MSE en este problema.

5. Estudie las gráficas de la función de densidad de probabilidad (PDF) del error para el MSE y MEE (mejor caso). Comente sobre las diferencies observadas justificando según la teoría. ¿Cuál es la relación entre el potencial de información y la esperanza de la PDF del error? ¿Qué implica minimizar o maximizar el potencial de información en términos de la PDF del error?

6. Considere ahora un modelo $Z(x) = c + ax + bx^2 + \nu$ y un regresor a $g(t; w) = w_0 + w_1 x + w_2 x^2$. Notar la inclusión de un término constante c y un peso w_0. ¿Por qué es necesario compensar el término constante (sesgo) cuando se usa MEE?

IMPORTANTE: El día Martes 13 de Octubre en horario de clases se realizará una sesión de laboratorio en la sala de computación del segundo piso del DIE donde se desarrollará la tarea. Se pide que al final de la sesión de laboratorio usted entregue un reporte de no más de dos planas con los resultados de su análisis y las discusiones pedidas en cada uno de los puntos de la tarea. La asistencia es OBLIGATORIA. La tarea es INDIVIDUAL.

IPYTHON NOTEBOOK: Para iniciar un servidor de *ipython notebook*, abra una consola de anaconda (Menú de inicio, anaconda, open anaconda shell) y escriba *ipython notebook*. Se debería abrir un browser en la dirección http://127.0.0.1:8888/. Una vez en la interfaz del *ipython notebook* utilice el explorador de carpetas y navegue hasta encontrar el archivo tarea2.ipynb adjunto. Se recomienda que guarde el archivo tarea2.ipynb en C:\\Anaconda\Examples. Una vez seleccionado el notebook se desplegará en el browser. Para ejecutar cada bloque de código presione SHIFT+ENTER, el notebook producirá los resultados y avanzará al siguiente bloque automáticamente. Puede seleccionar manualmente un bloque y volver a correrlo (SHIFT+ENTER) si lo desea.