UNIVERSIDAD ANÁHUAC QUERÉTARO

EXAMEN INTERSEMESTRAL

MATERIA: Aprendizaje de Máquina

CARRERA: **Ingeniería Biomédica para la Dirección. Ingeniería en Informática y Negocios Digitales. Ingeniería Mecatrónica.**

CLAVE: SIS-4309

SEMESTRE: 5°

PROFESOR: Jonathan Domínguez Aldana

FECHA: 08/10/2019

Nombre del Estudiante: ____

ID del estudiante: ____

INSTRUCCIONES

1. INSTRUCCIONES DE PUNTAJE: El examen consta de 100 puntos distribuidos en diferentes secciones. El valor de cada sección, así como sus instrucciones, están indicadas al inicio de la misma.

2. INSTRUCCIONES DE RECURSOS PERMITIDOS: No está permitido el uso de gadgets y la intercomunicación vía redes sociales, teléfono, mensajería o correo en el examen teórico. Esta actividad se considerará como un acto de deshonestidad académica.

3. INSTRUCCIONES DE DURACIÓN: Tiempo máximo para resolver el examen teórico: 200 minutos. Tiempo máximo para resolver el examen práctico: 1 semana.

4. INSTRUCCIONES DE FORMA: Lee cuidadosamente antes de contestar. Escribe tus respuestas en orden y con limpieza en las hojas anexas. Escribe con letra legible y sin faltas de ortografía. Recuerda que este instrumento valida tus conocimientos, por lo que se requiere de honestidad, disciplina y responsabilidad para contestarlo. Utiliza todas las herramientas conceptuales vistas hasta el momento en la clase. El examen es de carácter individual.

¡Éxito en su evaluación!

Sección teórica

1. Obtenga la **expresión vectorial** del vector de pesos $\boldsymbol{w}$ que **minimice** la función de costo $\boldsymbol{J}(\boldsymbol{w})$**,** representada como el error cuadrático medio **con regularización L2.** Donde $\boldsymbol{\lambda} > \mathbf{0}$ es la **constante** de **regularización**:

$$J(w) = \frac{1}{2}\sum_{i=1}^{m}\{y^{(i)} - w^T\phi(x^{(i)})\}^2 + \frac{\lambda}{2}w^T w$$

recordando que:

$$\vec{y} = \begin{bmatrix} y^{(1)} \\ y^{(2)} \\ \vdots \\ y^{(m)} \end{bmatrix} \qquad x^{(i)} = \begin{bmatrix} 1 \\ x_1 \\ \vdots \\ x_n \end{bmatrix} \qquad w = \begin{bmatrix} w_0 \\ w_2 \\ \vdots \\ w_k \end{bmatrix} \qquad \boldsymbol{\phi_j} = \boldsymbol{\phi}\left(\boldsymbol{x}^{(i)}\right)^{\boldsymbol{T}}$$

$$\boldsymbol{\phi}(\boldsymbol{x}) = \begin{bmatrix} \boldsymbol{\phi_0}(\boldsymbol{x}^{(1)}) & \cdots & \boldsymbol{\phi_{k-1}}(\boldsymbol{x}^{(1)}) \\ \vdots & \ddots & \vdots \\ \boldsymbol{\phi_0}(\boldsymbol{x}^{(m)}) & \cdots & \boldsymbol{\phi_{k-1}}(\boldsymbol{x}^{(m)}) \end{bmatrix}$$

2. Se sabe que la función tangente hiperbólica y la función sigmoide se encuentran representadas de la siguiente forma:

$$tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

$$\sigma(x) = \frac{1}{1 + e^{-x}}$$

a) Demuestre que la función tangente hiperbólica $\boldsymbol{tanh}$ y la función sigmoide $\boldsymbol{\sigma}$ están relacionadas por la siguiente expresión:

$$tanh(a) = 2\sigma(2a) - 1$$

b) Con el resultado del inciso a) demuestre que una combinación lineal general de funciones sigmoides representadas por:

$$y(x, w) = w_0 + \sum_{j=1}^{M} w_j \, \sigma\left(\frac{x - \mu_j}{s}\right)$$

y es equivalente a la combinación lineal de funciones ***tanh*** de la forma:

$$y(x, u) = u_0 + \sum_{j=1}^{M} u_j \, tanh\left(\frac{x - \mu_j}{2s}\right)$$

Encuentre las expresiones que relacionen los nuevos parámetros u a los parámetros originales w.

3. Considere un conjunto de datos, en donde cada variable de salida $\boldsymbol{y}^{(i)}$ se encuentra ponderada por un factor $\boldsymbol{r}^{(i)} > \mathbf{0}$, por lo que la función de costo estaría dada por:

$$J(w) = \frac{1}{2} \sum_{i=1}^{m} r^{(i)} \{y^{(i)} - w^T \phi(x^{(i)})\}^2$$

Encuentre una **expresión vectorial** para la solución del vector de pesos **w** que **minimice** la función de costo.

Como **ayuda**, se presentan los **siguientes cambios** de **variable** de **manera** que se **pueda llegar** a una **representación vectorial**:

$$\boldsymbol{\phi}'(\boldsymbol{x}^{(i)}) = \sqrt{\boldsymbol{r}^{(i)}} \, \boldsymbol{\phi}(\boldsymbol{x}^{(i)})$$

$$\boldsymbol{y}'^{(i)} = \sqrt{\boldsymbol{r}^{(i)}} \, \boldsymbol{y}^{(i)}$$

Sección práctica

Planteamiento del problema

Algunos dicen que el cambio climático es la mayor amenaza de nuestra época, mientras que otros dicen que es un mito basado en ciencia no fundamentada. Por ejemplo, el presidente de los Estados Unidos de Norteamérica, Donald Trump, cree firmemente que el cambio climático es una farsa.

Para comprobar cual de las dos posturas es acertada, se obtuvo un conjunto de datos recolectados por la institución Berkely Earth. El estudio, denominado "*Berkeley Earth Surface Temperature Study*", combinó 1,6 billones de reportes de temperatura de 16 archivos preexistentes.

La plataforma de competencias Kaggle publicó los datos recolectados por Berkeley Earth, los cuales consisten cinco archivos ".csv" distintos.

Para acceder a ellos, es necesario ingresar al siguiente link:
https://www.kaggle.com/berkeleyearth/climate-change-earth-surface-temperature-data

Tu labor, como ingeniero de inteligencia artificial, será analizar solamente los datos de temperatura de México, contenidos en el archivo "GlobalLandTemperaturesByCountry.csv".

De manera específica, el Congreso de la Unión de los Estados Unidos Mexicanos, te ha llamado como experto para demostrarle al presidente de México que si existe un incremento de temperatura en el país; de tal manera, que no se deje convencer por el presidente norteamericano.

Tus tareas, para convencer al presidente Andrés Manuel López Obrador del cambio climático, son las siguientes:

1. **Extracción y preparación de datos:**

   Extrae los datos del archivo separado por comas "GlobalLandTemperaturesByCountry.csv", calcula los promedios anuales de temperatura desde el año 1752 al año 2015 y grafica las temperaturas globales. El resultado obtenido debe ser igual al representado por la Figura 1.

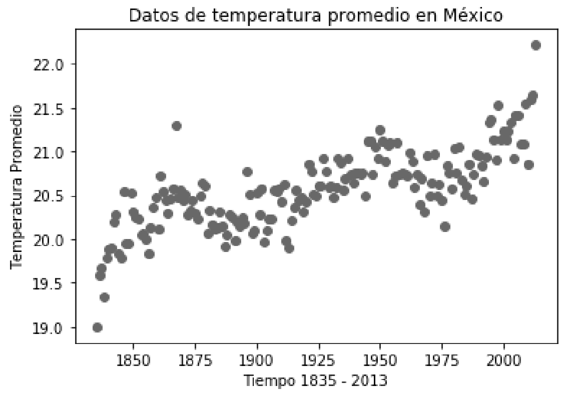

Figura 1. Gráfica de la temperatura promedio anual en México.

2. Ecuaciones normales – Funciones de base polinómicas:

a) Implemente las funciones de base polinómicas expresadas como:

$$\phi_j(x) = \left(x^{(i)}\right)^j$$

$$\phi(x) = \begin{bmatrix} 1 & {x^{(1)}}^T & \cdots & \left({x^{(1)}}^T\right)^{(k-1)} \\ \vdots & \vdots & & \vdots \\ 1 & {x^{(m)}}^T & \cdots & \left({x^{(m)}}^T\right)^{(k-1)} \end{bmatrix}$$

b) Obtenga el mejor vector de pesos w que minimice el error de cuadrático medio con regularización L2. Es decir, obtenga e implemente las ecuaciones normales con regularización L2.

c) Ajustar los datos de temperatura de entrenamiento y validación con polinomios de grado $j = 1, 2, \ldots, 10$. Para cada polinomio probar al menos cinco diferentes valores de la constante de regularización λ. (Se obtendrán 50 modelos distintos).

d) Graficar y/o guardar cada una de las curvas ajustadas (en este caso 50) para los datos de entrenamiento y validación, junto con su respectivo errores. La Figura 2 muestra un ejemplo:

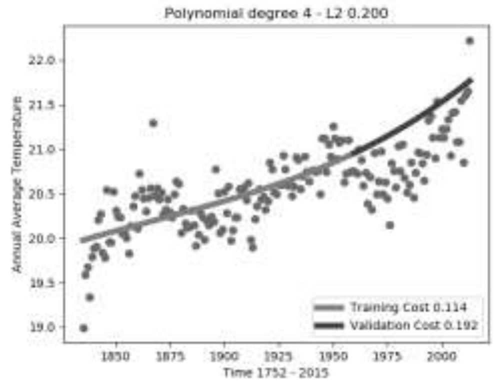

Figura 2. Modelo polinómico de grado 4 con una constante $\lambda = 0.2$ ajustado a los datos de temperaturas globales.

e) Seleccionar el mejor modelo con capacidad de generalización, de los modelos del inciso c), y utilizarlo para predecir las temperaturas promedio anuales en México desde el año 2016 hasta el año 2100. La Figura 3 muestra un ejemplo:

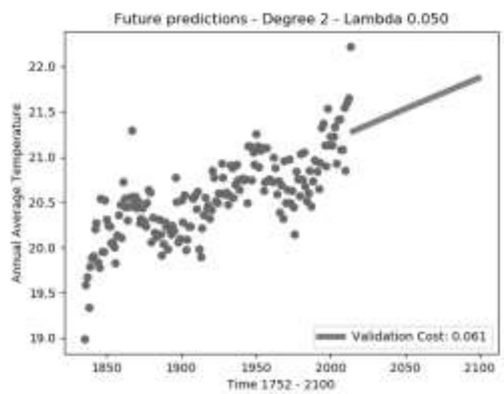

Figura 3. Predicciones de temperatura del año 2016 hasta el año 2100 usando el mejor modelo obtenido.

3. Descenso por gradiente– Funciones de base polinómicas:

Incisos a), b), c) y d):

Repite los incisos 2b, 2c, 2d y 2e, pero ajustando los modelos con descenso por gradiente usando funciones de base polinómicas.

4. Descenso por gradiente– Funciones de base gaussiana:

a) Implemente las funciones de base gaussiana expresadas como:

$$\phi_j(x) = e^{\left(-\frac{(x-\mu_j)^2}{2\sigma^2}\right)^j}$$

$$\phi(x) = \begin{bmatrix} e^{\left(-\frac{(x^{(1)}-\mu_1)^2}{2\sigma^2}\right)^T} & \cdots & e^{\left(-\frac{(x^{(m)}-\mu_1)^2}{2\sigma^2}\right)^T} \\ \vdots & \ddots & \vdots \\ e^{\left(-\frac{(x^{(1)}-\mu_k)^2}{2\sigma^2}\right)^T} & \cdots & e^{\left(-\frac{(x^{(m)}-\mu_k)^2}{2\sigma^2}\right)^T} \end{bmatrix}$$

donde σ^2 es constante y el parámetro que varía es μ_j, es decir se implementan k gaussianas cada una con diferente μ_j.

b) Obtenga el mejor vector de pesos w que minimice el error de cuadrático medio con regularización L2. Es decir, obtenga e implemente el descenso por gradiente con regularización L2.

c) Ajustar los datos de temperatura de entrenamiento y validación con $k = 5, 6, \ldots, 10$ gaussianas, cada gaussiana con una μ_j diferente y una σ^2 constante (Total de 5 modelos).

Para cada modelo asignar cinco valores diferentes de σ^2 y cinco valores diferentes de la constante de regularización λ (Total de 125 modelos).

d) Graficar y/o guardar cada una de las curvas ajustadas (en este caso 125) para los datos de entrenamiento y validación, junto con su respectivo errores.

e) Seleccionar el mejor modelo con capacidad de generalización, de los modelos del inciso c), y utilizarlo para predecir las temperaturas promedio anuales en México desde el año 2016 hasta el año 2100.