**Documento de Pruebas realizadas**

**Introducción:**

Dentro del problema de la estimación de la altura de la ola en la cercanía de la costa a partir de un modelo de predicción de oleaje de alta mar, se evaluaron tres algoritmos de regresión. Regresión Lineal, Redes Neuronales, Maquina de soporte de vectores.

A continuación se detallan los datos de entrenamiento utilizados, la perfomance obtenida con cada algoritmo y una comparativa entre estos.

**Conjunto de entrenamiento:**

**Atributos:**

* Altura de la ola (wave height) (metros): Es una medida de la altura de la ola y se corresponde a lo que un observador entrenado considerario la altura promedio.
* Dirección de la ola(grados): la dirección promedio de las olas.
* Periodo de la ola (seg.):el tiempo entre olas.
* Observación visual(metros)(Atributo de clase): Altura de la ola al momento de romper medida desde la costa por un observador experto.

Existe una observación visual por día correspondiente a la observación de la ola más alta registrada durante el día.

Los atributos provenientes de la salida del modelo WW3 están disponibles 8 veces al día.

El proceso para acoplar la observación visual diaria con una de las 8 salidas del pronóstico del ww3 fue el siguiente:

1. Eliminar todas las salidas del ww3 que no pertenecen al horario de luz de la costa de Hawái.
2. Elegir la salida del ww3 con mayor altura de ola correspondiente al día de la observación visual.

El razonamiento para 1 es que las observaciones visuales fueron hechas durante el periodo de luz de la costa de hawaii, por lo tanto no tiene sentido utilizar los datos del WW3 que pertenecen a la franja horaria nocturna.

El razonamiento para 2 es que como la observación visual pertenece a la altura de ola más alta registrada durante el día entonces la salida del ww3 asociada con esa lectura debería ser también la más alta registrada en el día.

**Prueba 1.(Regresión Lineal)**

**Periodo:** 1 Marzo del 2002 al 31 de Diciembre del 2002

**Grilla del ww3**:Latitud: 22° 0'0.00"N Longitud: 158°46'7.62"O

**Lugar de la observación visual:** North Shore, Oahu, Sunset.

**Número total de instancias:** 364.

**Filtrados:** Ninguno.

**Algoritmo:** weka.classifiers.functions.LinearRegression -S 0 -R 1.0E-8

**Resultados**:

Coeficiente de correlación 0.8175

Error absoluto medio 1.1041.

Error absoluto cuadrático medio 1.5995

La interpretación del error absoluto es que existe un error promedio de +/- 1,10 metros en las predicciones de la altura de la ola.

Mientras que el error cuadrado medio que penaliza las desviaciones mas grandes obtenemos un error promedio de +/-1.6 metros.

El siguiente grafico corresponde al desempeño del clasificador obtenido utilizando un conjunto de pruebas que nunca fue utilizado para entrenamiento:

**Prueba 2.(Red Neuronal)**

**Periodo:** 1 Marzo del 2002 al 31 de Diciembre del 2002

**Grilla del ww3**:Latitud: 22° 0'0.00"N Longitud: 158°46'7.62"O

**Lugar de la observación visual:** North Shore, Oahu, Sunset.

**Número total de instancias:** 364.

**Filtrados:**Ninguno.

**Algoritmo:**

weka.classifiers.functions.MultilayerPerceptron -L 0.99 -M 0.01 -N 1000 -V 0 -S 3 -E 20 -H t –D

**Comentarios:**

Se evaluó este algoritmo utilizando diferentes parámetros de configuración. Esta fue la configuración que mejores resultados obtuvo:

*Taza de aprendizaje:* 0.99

*Numero de capaz ocultas:* 4(Atributos+ clases).

*Decaimiento de la taza de aprendizaje:* activado.

*Training Rate(Epochs):* 1000

*Momentum:* 0.01

*Normalización de atributos*= activada

*Reset*= activado

**Resultados**:

Coeficiente de correlación 0.8417

Error absoluto medio 1.0346.

Error absoluto cuadrático medio 1.4996

La interpretación del error absoluto es que existe un error promedio de +/- 1 metro en las predicciones de la altura de la ola.

Mientras que el error cuadrado medio que penaliza las desviaciones mas grandes obtenemos un error promedio de +/-1.5 metros.

El siguiente grafico corresponde al desempeño del clasificador obtenido utilizando un conjunto de pruebas que nunca fue utilizado para entrenamiento:

**Prueba 3.(Maquina de Soporte Vectorial)**

**Periodo:** 1 Marzo del 2002 al 31 de Diciembre del 2002

**Grilla del ww3**:Latitud: 22° 0'0.00"N Longitud: 158°46'7.62"O

**Lugar de la observación visual:** North Shore, Oahu, Sunset.

**Número total de instancias:** 364.

**Filtrados:**Ninguno.

**Algoritmo:**

weka.classifiers.functions.SMOreg -S 0.0080 -C 0.7 -E 2.0 -G 0.5 -A 23444 -T 0.0080 -P 1.0E-12 -N 0 –R

**Comentarios:**

Se evaluó este algoritmo utilizando diferentes parámetros de configuración y se muestran los mejores resultados obtenidos.

**Resultados**:

Coeficiente de correlación 0.8476

Error absoluto medio 0.9795

Error absoluto cuadrático medio 1.5363

La interpretación del error absoluto es que existe un error promedio de +/- 1 metro en las predicciones de la altura de la ola.

Mientras que el error cuadrado medio que penaliza las desviaciones mas grandes obtenemos un error promedio de +/-1.5 metros.

El siguiente grafico corresponde al desempeño del clasificador obtenido utilizando un conjunto de pruebas que nunca fue utilizado para entrenamiento:

**Comparación de los algoritmos de regresión:**

Analizando el siguiente cuadro se puede ver que no existe una gran diferencia en perfomance entre los diferentes algoritmos. El algoritmo con peor desempeño ha sido el de regresión lineal, siendo superado en todos los indicadores por los algoritmos de red neuronal y de soporte vectorial. Entre estos dos últimos el algoritmo de soporte vectorial ofrece un aumento ínfimo en la correlación como así también en el error medio absoluto, no siendo así en el error cuadrático medio, donde la red neuronal posee una pequeña ventaja.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Regresión Lineal** | **Red Neuronal** | **Soporte Vectorial** |
| **Correlación** | 0.8175 | 0.8417 | 0.8476 |
| **Error Medio Absoluto** | 1.1041 | 1.0346 | 0.9795 |
| **Error Cuadrático Medio Absoluto** | 1.5995 | 1.4996 | 1.5363 |

**A continuación se grafican los valores de altura de ola predichos por los algoritmos evaluados junto a la observación visual correspondiente.**