

Funciones de Base Radial (RBF)

García Prado, Sergio

`sergio@garciparedes.me`

29 de mayo de 2017

Resumen

En este documento se describen las redes neuronales formada por funciones de base radial. Además, se realiza un experimento sobre el conjunto de datos Wine para comprobar el número de clusters optimo para este conjunto de datos. Posteriormente se utilizan las funciones de base radial en combinación con un perceptrón multicapa. Finalmente se discuten los resultados de los distintos experimentos.

1. INTRODUCCIÓN

1.1. FUNCIONES DE BASE RADIAL

Las redes neuronales con funciones de base radial son una estrategia de aprendizaje automático basada en medidas de distancia como valor de bondad. Utilizan una estructura básica de 3 capas (aunque se puede ampliar a más capas) siendo la primera (capa de entrada) una capa sin procesado compuesta por una neurona por atributo de entrada.

La capa oculta está formada por funciones de base radial, que indican la medida de cercanía de las instancias a un determinado punto. La capa de salida está formada por funciones de activación. El aprendizaje de este tipo de redes se basa en la corrección de los pesos de las funciones de activación en la capa de salida así como la modificación de los centros en las funciones de base radial de la capa oculta.

Existen distintas funciones de base radial que se pueden utilizar en este tipo de redes neuronales. Sin embargo, la característica principal que deben poseer es la de maximizar su valor conforme la instancia a examinar se asemeje al centro del cluster.

2. EXPERIMENTOS SOBRE WINE DATASET

El conjunto de datos utilizado para los experimentos que se describirán en esta sección es **Wine Data Set** [UCI], compuesto por **178 instancias**. Dichas instancias contienen **12 atributos** de carácter numérico. La **clase de destino es de carácter numérico**. Sin embargo, en este caso es de tipo entero, por lo que puede ser vista como una variable categórica. En cuanto a la metodología experimental, se ha seguido una estrategia de **HoldOut** con particionamiento de los datos de manera que $\frac{2}{3}$ son utilizados en la fase de entrenamiento y $\frac{1}{3}$ en la de test.

El objetivo de este experimento es la búsqueda del número óptimo de clusters o centros a utilizar sobre el conjunto de datos *Wine* para la estrategia de aprendizaje automático basada en *Funciones de Base Radial (RBF)*. Una vez obtenido el valor óptimo se utilizará dicha red (la capa oculta) como preprocesamiento de la entrada para después clasificar las instancias mediante un *Perceptrón Multicapa* con 10 neuronas en la capa oculta.

# Clusters	Tasa de Error
3	0.72131147541
4	0.0491803278689
5	0.16393442623
6	0.016393442623
7	0.0491803278689
8	0.0491803278689
9	0.0491803278689
10	0.0327868852459
11	0.0327868852459
12	0.0327868852459
13	0.0327868852459
14	0.0327868852459
15	0.0491803278689

Tabla 1: Tasas de error obtenidas para el experimento de Holdout sobre el conjunto de datos Wine mediante el uso de Funciones de Base Radial (RBF) con distinto número de clusters

Los resultados obtenidos durante la fase de búsqueda del número óptimo de centros o clusters para la red compuesta por funciones de base radial se muestra de manera tabular en la tabla 1. También se ha incluido una representación gráfica de dichos resultados para facilitar su comparación, esto se muestra en la figura 1. Tal y como se puede apreciar, el valor óptimo es **6 centros**. Algo a destacar es el caso de 6 centros, que obtiene tasas de error muy elevadas. La razón puede ser debida a que el número de clases de destino también está formada por 3 categorías por lo cual no tiene centros de holgura para realizar combinaciones de los mismos en la capa de salida. En cuanto a los resultados mayores que 3, en el resto de casos los resultados son muy aceptables, cercanos al óptimo encontrado (6 centros).

El siguiente paso ha sido utilizar dicho valor para crear una estructura de red neuronal compuesta de 2 capas ocultas. La primera de ellas utiliza funciones de base radial mientras que la segunda está formada por funciones no lineales con función de activación (por ser un perceptrón multicapa). Por tanto, la estructura de la red es **12-6-10-1** ya que se ha utilizado.

La razón de que la capa de entrada esté compuesta por 12 neuronas se debe a que es el número de atributos de entrada, la primera capa oculta está compuesta por 6 neuronas con función de base radial, la segunda capa oculta por 10 neuronas de tipo perceptrón y la capa de salida está formada por una única neurona. Por lo tanto, el modelo actúa como una regresión. La razón de que el experimento se haya realizado así es poder calcular la tasa de error mediante el error relativo de manera manual. Para ello se ha utilizado un script escrito en Matlab al cual se puede acceder a través de <https://gist.github.com/garciparedes/01f6c49054d26a8122e96f349f59b146>

La construcción se ha realizado utilizando la suite de aprendizaje automático *WEKA* mediante el apoyo en el filtro de entrada *ClusterMembership*(<http://weka.sourceforge.net/doc.stable/>

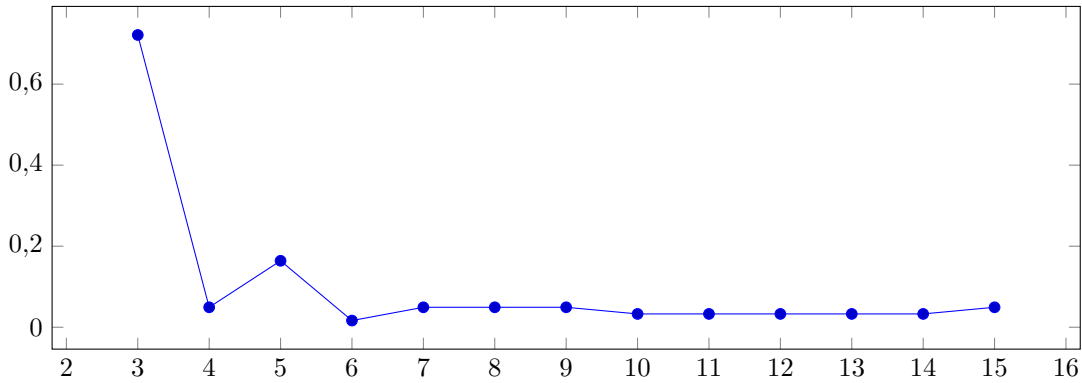


Figura 1: Tasas de error obtenidas para el experimento de Holdout sobre el conjunto de datos Wine mediante el uso de Funciones de Base Radial (RBF) con distinto número de clusters

RBF K -centros + MLP 10	
Datos	Tasa de Error ($K =$)
	6
Pruebae	0,016393442623

Tabla 2: Tasa de error obtenida para el experimento de Holdout sobre el conjunto de datos Wine mediante el uso de Funciones de Base Radial (RBF) combinadas con un Perceptrón Multicapa con 10 neuronas en la capa oculta

`weka/filters/unsupervised/attribute/ClusterMembership.html`), que indica el grado de pertenencia de los atributos de entrada a los centros prefijados, de manera que actúa de manera muy similar al algoritmo de aprendizaje no supervisado **k-means**.

Los resultados obtenidos tras la realización del experimento descrito se muestran en la tabla 2. Tal y como se puede apreciar, la tasa de error es igual a la obtenida únicamente mediante el uso de funciones de base radial (no hay mejora), por tanto, se cree que una implementación más sencilla obviando el uso del Perceptrón Multicapa es una solución más acertada en este caso. Además, las tasas de error obtenidas mediante el uso de este son muy bajas.

REFERENCIAS

- [CCAG17] Teodoro Calonge Cano and Carlos Javier Alonso González. Técnicas de Aprendizaje Automático, 2016/17.
- [GP17] Sergio García Prado. Funciones de base radial. <https://github.com/garciparedes/machine-learning-radial-basis-functions>, 2017.
- [too] Weka. <http://www.cs.waikato.ac.nz/ml/weka/>.
- [UCI] UCI Machine Learning Repository. Wine Data Set. <http://archive.ics.uci.edu/ml/datasets/Wine>.