

Trabajo práctico 2. Gatos vs. Perros



El objetivo de este trabajo práctico es construir un clasificador de imágenes de animales que corresponden a 2 clases: gatos y perros. Las imágenes se obtuvieron de diversos bancos de imágenes, competencias¹ y sitios de Internet.

Se entregará:

1. Un dataset de entrenamiento con imágenes de gatos y perros. Las imágenes son archivos .jpg. Será responsabilidad de cada grupo tomar esos objetos, codificarlos y extraerles las características que consideren adecuadas para ser incorporados en los procesos de aprendizaje. Además se proveerá un conjunto pequeño de imágenes, de la misma pinta que recibirán el día de la competencia para oficiar de validación del software.
2. El día de la competencia los grupos recibirán un dataset con imágenes semejante al de entrenamiento salvo que no incluye la clase en el nombre de los archivos. Los docentes retendremos las clases originales correspondientes a estos casos hasta el momento de evaluar la competencia.

Se pide:

¹ <https://www.kaggle.com/c/dogs-vs-cats>

1. Elaborar un modelo de reconocimiento de las imágenes a partir del dataset entregado basado en técnicas de aprendizaje supervisado.
2. El modelo deberá estar compuesto por lo menos por 3 métodos diferentes de aprendizaje ensamblados. Pueden utilizar cualquiera de los métodos presentados en la materia y soportados por los softwares que hemos utilizado: scikit-learn, R, Weka, etc. Pueden usar cualquier otra opción pero debe estar previamente consensuada con los docentes.
3. El día de la entrega, los grupos deberán entregar el software para clasificar las imágenes junto con el informe. Los programas deberán poder ejecutarse en las PCs del laboratorio. Se proveerá un conjunto de
4. El día de la competencia contarán con 30 minutos para procesar el dataset de competencia y devolver un archivo "grupoNN.txt" donde cada fila tendrá: "Id, Clase" siendo Id el nombre del archivo y Clase.
5. Pasado el tiempo establecido se procesarán los archivos entregados por cada grupo y se confeccionará el ranking para obtener la posición obtenida.
6. Deberán tomar los recaudos para tener instalados los modelos desarrollados en las máquinas del laboratorio o en sus propias máquinas de tal manera de poder participar en la competencia a la hora establecida.
7. Al final de la competencia, deberán realizar una breve exposición ante los docentes y demás grupos, de las principales características del trabajo realizado.
8. Producir y entregar un informe que contenga, como mínimo: presentación sintética de la solución propuesta y resultado de las pruebas realizadas que justifican la elección; estructura, contenido y justificación de las características extraídas del dataset; descripción general de las tareas realizadas y problemas encontrados; variantes utilizadas con las respectivas justificaciones de adopción o rechazo; exposición detallada de la solución propuesta. El documento a entregar deberá cumplir con los requisitos normales de un "informe profesional" y debe incluir como mínimo: resumen ejecutivo, objetivo, arquitectura de la solución, pruebas de eficiencia y análisis de resultados, conclusiones, ideas para futuros desarrollos, bibliografía y referencias.

Aclaraciones:

1. El presente trabajo deberá elaborarse en grupos de 2 personas.
2. La **fecha límite de entrega** del trabajo es el martes 24/11/2015 a las 23:59.
3. La **competencia** será el día miércoles 2/12/2015 a las 13:00 en el laboratorio.
4. Se permitirá entregar el trabajo sin presentarse en la competencia solo a quienes lo avisen con una semana de anticipación. La aceptación de éstos quedará a exclusivo criterio de los docentes.
5. Se podrán pedir pruebas de integridad y autoría, es decir, verificar que la salida solicitada es producto del modelo presentado y que el modelo fue construido según lo requerido en el presente documento.
6. Los procesos de codificación y extracción de características (atributos) deberá estar debidamente documentado y referenciado.

7. La **evaluación** será grupal y se basará en la calidad del informe; la originalidad, practicidad y coherencia técnica de la solución; la corrección y solidez de las pruebas realizadas. Todos los integrantes del grupo obtendrán la misma calificación. La posición del grupo en el ranking de la competencia **no** forma parte de la evaluación.
8. El presente es un trabajo de evaluación para una materia de Aprendizaje Automático, por lo tanto, deben aplicarse las prácticas aceptadas en esta especialidad. No se aceptarán soluciones de otro tipo, por eficientes que sean.