Universitat Autònoma de Barcelona

2023/2024

Aprendizaje Automático para Visión por Computador

Código: 44774
Créditos ECTS: 6

Titulación	Tipo	Curso	Semestre
4318299 Visión por Computador	OB	0	1

Contacto

Nombre: Maria Isabel Vanrell Martorell

Correo electrónico: maria.vanrell@uab.cat

Idiomas de los grupos

Puede consutarlo a través de este enlace. Para consultar el idioma necesitará introducir el CÓDIGO de la asignatura. Tenga en cuenta que la información es provisional hasta el 30 de noviembre del 2023.

Equipo docente

Ramon Baldrich Caselles

Fernando Luis Vilariño Freire

Dimosthenis Karatzas

Pau Rodriguez Lopez

Guillem Arias Bedmar

Luis Gomez Bigorda

Prerrequisitos

- Un grado en Ingeniería, Matemáticas, Física o similar.
- Competencias de programación en Python.

Objetivos y contextualización

Coordinador del módulo: Dr. Ramon Baldrich Caselles
El objetivo de este módulo es presentar las técnicas de aprendizaje automático para resolver problemas de visión por computador. El aprendizaje automático se ocupa del análisis automático de datos a gran escala. Hoy en día, conforma los conceptos básicos de muchos métodos de visión por computadora, especialmente aquellos relacionados con el reconocimiento o la clasificación de patrones visuales, donde los 'patrones' abarcan imágenes de objetos del mundo, escenas y secuencias de video de acciones humanas, por nombrar algunos.

Este módulo presenta los fundamentos y las técnicas más importantes para la clasificación de patrones

visuales, centrándose principalmente en métodos supervisados. Además, se abordan temas relacionados como descriptores de imágenes y reducción de dimensionalidad. En la medida de lo posible, todas estas técnicas se prueban y evalúan en un proyecto práctico relacionado con la descripción de la escena a partir de imágenes, junto con las métricas y los procedimientos estándar para la evaluación del rendimiento, como las curvas de recuperación de precisión y la validación cruzada en k veces.

Los resultados del aprendizaje son:

(a) Distinga los principales tipos de técnicas de ML para la visión por computadora: supervisado versus no supervisado, generativo versus discriminativo, espacio de características originales versus kernelización de vectores de características.

(b) Conozca los puntos fuertes y débiles de los diferentes métodos, en parte aprendidos mientras resuelve un problema real de clasificación de patrones.

(c) Ser capaz de usar implementaciones de métodos existentes y construirlas desde cero.

El módulo profundiza en dos enfoques principales para introducir ML en el problema de clasificación de imágenes. Usando: a) descripción de imagen hecha a mano, b) descripción de imagen basada en datos. En el primer caso se usa Bag of Words, en el segundo el enfoque de aprendizaje profundo. El contenido de DL se desarrolla, proporcionando tanto bases teóricas de las diferentes partes de los modernos sistemas de Neural Networs como las mejores prácticas para aplicarlo en aplicaciones reales.

Resultados de aprendizaje

- CA06 (Competencia) Conseguir los objetivos de un proyecto de visión realizado en equipo.
- KA03 (Conocimiento) Identificar qué métodos de aprendizaje computacional se pueden usar en función de los datos para la resolución de un problema de visión.
- KA10 (Conocimiento) Seleccionar los mejores procedimientos de experimentación que se deben diseñar para el aprendizaje computacional desde el entrenamiento hasta la evaluación.
- KA16 (Conocimiento) Reconocer las dimensiones de género, medioambientales y éticas de los sistemas de visión y su aplicación.
- SA03 (Habilidad) Aplicar y evaluar técnicas de aprendizaje computacional para solucionar un problema particular.
- SA13 (Habilidad) Calcular la huella de carbono de cualquier experimento que requiera entrenar una red neuronal profunda.
- SA14 (Habilidad) Detectar sesgos en conjuntos de datos de aprendizaje que permitan evitar la construcción de sistemas que discriminen socialmente.
- SA17 (Habilidad) Preparar presentaciones orales que permitan debatir los resultados del desarrollo de un proyecto de visión.

Contenido

1. Introducción al aprendizaje computacional
2. Entorno experimental
3. Codificaciones: SVM y Random Forest
4. Introducción a las Redes Neuronales
5. Introducción al Aprendizaje Profundo (Deep Learning, DL)
6. Redes Neuronales Convolucionales (CNN)
7. Entrenamiento: preprocesado de datos, inicialización, optimización del gradiente
8. Clasificación de imágenes
9. Entender y visualización de CNNs
10. Métodos eficientes para el Aprendizaje Profundo

Metodología

Sesiones supevisadas: *(Algunes d'aquestes sesiones podrien ser en línea síncronas)*

- Sesiones magistrales, donde los profesores explicarán contenidos generales de los diferentes temas. La mayoria seran necesarios para la resolución de problemas.

Sesiones dirigidas:

- Sesiones de proyecto, donde los objetivos y problemas de los proyectos se presentarán y discutirán. Los estudiantes deberan interactuar con el coordinador de proyecto sobre los problemas surgidos y las ideas aportadas para resorverlos. (Approx. 1 hora/semana)
- Sesiones de presentación, donde los estudiantes harán una presentación oral sobre como han solucionado el probema y sobre los resultados obtenidos.
- Sesión de exámen, donde los estudiantes son evaluados individualmente, demostrando la adquisición de los conocimientos desarrollados y la capacidad de resolución de problemas asociados.

Trabajo autónomo:

- Los estudiantes deberan estudiar y trabajar autonomamente con los materiales derivados de las clases magistrales y de las sesiones de proyecto.
- Los estudiantes trabajaran en grupo para resolver los problemas planlteados en los proyectos con los siguientes entregables:
    - Codigo
    - Informe
    - Presentación oral

Nota: se reservarán 15 minutos de una clase dentro del calendario establecido por el centro o por la titulación para que el alumnado rellene las encuestas de evaluación de la actuación del profesorado y de evaluación de la asignatura o módulo.

Actividades

Título	Horas	ECTS	Resultados de aprendizaje
Tipo: Dirigidas			
Sesiones teóricas	20	0,8	KA03, KA10, KA16, KA03
Tipo: Supervisadas			
Sesiones de seguimiento de proyecto	8	0,32	CA06, SA03, SA13, SA14, SA17, CA06
Tipo: Autónomas			
Trabajo autónomo	113	4,52	CA06, SA03, SA13, SA14, SA17, CA06

Evaluación

La Nota final de este módulo se calcula con la seguiente fórmula:

Nota final = 0.4 x Examen + 0.55 x Proyecto+ 0.05 x Asistencia

donde,

Examen: es la nota obtenida en el examen final del módulo (tiene que ser >= 3).

Assistencia: es la nota derivada de la asistencia a las sesiones (mínimo 70%)

Proyecto: es la nota que pone el coordinador del proyecto de acuerdo con los resultados del seguimiento semanal del proyecto y de las entregas (tiene que ser >= 5). Todo de acuerdo con criterios específicos, tales como:

- Participación en les sesiones de discusión y en el trabajo en equipo (evaluación entre-miembros)
- Entrega de las partes obligatorias y opcionales del proyecto.
- Desarrollo del código (estilo, comentarios, etc.)
- Informe (justificación de las decisiones tomadas en el desarrollo del proyecto)
- Presentación (Presentación y demostración del proyecto desarrollado)

Sólo los estudiantes que han suspendido (nota final < 5.0) podran hacer el examen de recuperación.

Actividades de evaluación continuada

Título	Peso	Horas	ECTS	Resultados de aprendizaje
Asistencia a sesiones	0.05	0,5	0,02	CA06, KA03, KA10, KA16
Examen	0.4	2,5	0,1	KA03, KA10, KA16, SA03
Proyecto	0.55	6	0,24	CA06, SA03, SA13, SA14, SA17

Bibliografía

Artículos en revistas:

1. Barber, D. "*Bayesian Reasoning and Machine Learning"*. Cambridge University Press, 2012.
2. Yoshua Bengio. "*Learning Deep Architectures for AI"*. Foundations and Trends in Machine Learning, Vol. 2, No. 1, 2009.
3. Christopher J. C. Burges. "*Dimension Reduction: A Guided Tour"*. Foundations and Trends in Machine Learning, Vol. 2, No. 4, 2009.
4. Christoph H. Lampert. "*Kernel Methods in Computer Vision"*. Foundations and Trends in Computer Graphics and Vision, Vol. 4, No. 3, 2008.
5. Tinne Tuytelaars and Krystian Mikolajczyk. "*Local Invariant Feature Detectors: A Survey"*. Foundations and Trends in Computer Graphics and Vision, Vol. 3, No. 3, 2007.

Libros:

1. Ian Goodfellow, Yoshua Bengio and Aaron Courville. "*Deep Learning*". 2016. Cambridge, MA, USA: The MIT Press. ISBN: 978-0262035613
2. Mehryar Mohri, Afshin Rostamizadeh, and Ameet Talwalkar, "*Foundations of Machine Learning"* MIT Press, 2012. http://www.cs.nyu.edu/~mohri/mlbook/
3. Z.H. Zhou. *Ensemble Methods: Foundations and Algorithms*. Chapman & Hall/CRC, 2012.

Informes técnicos:

1. Criminisi, A. and Shotton, J. and Konukoglu, E. "*Decision Forests for Classification, Regression, Density Estimation, Manifold Learning and Semi-Supervised Learning"*. Technical report MSR-TR-2011-114. Microsoft Research, 2011. http://research.microsoft.com/pubs/155552/decisionForests_MSR_TR_2011_114.pdf

Software

Entorno de programación en Python con especial atención a las librerías de visión por computador y Keras