

**Aprenentatge Computacional per a la Visió per
Computador**

Codi: 43083

Crèdits: 6

Titulació	Tipus	Curs	Semestre
4314099 Visió per Computador / Computer Vision	OB	0	1

Professor/a de contacte

Nom: Ramón Baldrich Caselles

Correu electrònic: Ramon.Baldrich@uab.cat

Equip docent

Ernest Valveny Llobet

Maria Vanrell Martorell

Juan Carlos Moure Lopez

Fernando Luis Vilariño Freire

Dimosthenis Karatzas

Pau Rodriguez Lopez

Lluís Gomez Bigorda

Utilització d'idiomes a l'assignatura

Llengua vehicular majoritària: anglès (eng)

Prerequisits

- Un grau en Enginyeria, Matemàtiques, Física o similar.
- Competències en programació en Python.

Objectius

L'objectiu d'aquest mòdul és introduir les tècniques d'aprenentatge automàtic per resoldre problemes de visió per ordinador. L'aprenentatge automàtic tracta de l'anàlisi automàtica de dades a gran escala. Avui en dia s'ajusta als conceptes bàsics de molts mètodes de visió per ordinador, especialment els relacionats amb el reconeixement visual o la classificació de patrons, on els "patrons" engloben imatges d'objectes mundials, escenes i seqüències de vídeo d'accions humanes, per citar-ne algunes.

Aquest mòdul presenta els fonaments i les tècniques més importants per a la classificació de patrons visuals, centrats principalment en mètodes supervisats. També es tracten temes relacionats com els descriptors d'imatges i la reducció de dimensionalitat. En la mesura del possible, totes aquestes tècniques es proven i s'avaluen en un projecte pràctic sobre descripció d'escenes a partir d'imatges, en funció de les mètriques i procediments estàndards d'avaluació de rendiment com ara corbes de record de precisió i validació creuada de k.

Els resultats d'aprenentatge són:

(a) Distingiu els principals tipus de tècniques de ML per a la visió per ordinador: supervisat vs no supervisat, generatiu vs. discriminatori, espai de característiques original i nucli de vectors de característiques.

(b) Conèixer els punts forts i febles dels diferents mètodes, en part après durant la resolució d'un problema de classificació de patrons reals.

(c) Ser capaç d'utilitzar implementacions de mètodes existents i construir-les des de zero.

El mòdul aprofundeix en dues aproximacions principals per introduir ML al problema de classificació d'imatges. Utilitzant: a) descripció de la imatge amb descriptors definits manualment, b) descripció de la imatge basat en dades. En el primer cas s'utilitza el Bag of Words, en el segon, l'enfocament de l'aprenentatge profund. El contingut de DL es desenvolupa àmpliament proporcionant tant les bases teòriques de les diferents parts de les modernes acrílites de Neural Networks, com les bones pràctiques per aplicar-lo en aplicacions reals.

Competències

- Assumir tasques de responsabilitat en la gestió de la informació i el coneixement.
- Comprendre, analitzar i sintetitzar els coneixements avançats que hi ha en l'àrea, així com proposar idees innovadores.
- Conceptualitzar alternatives de solucions complexes per a problemes de visió i crear prototips que demostrin la validesa del sistema proposat.
- Identificar els conceptes i aplicar les tècniques fonamentals més adequades per resoldre els problemes bàsics de la visió per computador.
- Planificar, desenvolupar, avaluar i gestionar solucions per a projectes en els diferents àmbits de la visió per computador.
- Que els estudiants sàpiguin aplicar els coneixements adquirits i la seva capacitat de resolució de problemes en entorns nous o poc coneguts dins de contextos més amplis (o multidisciplinaris) relacionats amb la seva àrea d'estudi.
- Que els estudiants tinguin les habilitats d'aprenentatge que els permetin continuar estudiant, en gran manera, amb treball autònom a autodirigit.
- Seleccionar les eines de programari i els conjunts d'entrenament més adequats per desenvolupar les solucions per als problemes de visió per computador.
- Tenir coneixements que aportin la base o l'oportunitat de ser originals en el desenvolupament o l'aplicació d'idees, sovint en un context de recerca.
- Treballar en equips multidisciplinaris.

Resultats d'aprenentatge

1. Assumir tasques de responsabilitat en la gestió de la informació i el coneixement.
2. Comprendre, analitzar i sintetitzar els coneixements avançats que hi ha en l'àrea, així com proposar idees innovadores.
3. Identificar els algorismes bàsics d'aprenentatge computacional i la seva aplicació.
4. Identificar les millors representacions que es puguin definir per resoldre problemes d'aprenentatge computacional
5. Que els estudiants sàpiguin aplicar els coneixements adquirits i la seva capacitat de resolució de problemes en entorns nous o poc coneguts dins de contextos més amplis (o multidisciplinaris) relacionats amb la seva àrea d'estudi.
6. Que els estudiants tinguin les habilitats d'aprenentatge que els permetin continuar estudiant, en gran manera, amb treball autònom a autodirigit.
7. Seleccionar tècniques d'aprenentatge computacional i entrenar-les perquè solucionin un projecte concret.
8. Tenir coneixements que aportin la base o l'oportunitat de ser originals en el desenvolupament o l'aplicació d'idees, sovint en un context de recerca.
9. Treballar en equips multidisciplinaris.
10. Utilitzar tècniques d'aprenentatge computacional per planificar, desenvolupar, avaluar i gestionar una solució per a un problema concret

Continguts

1. Introducció a l'aprenentatge computacional
2. Entorn experimental
3. Codificacions: SVM i Random Forest
4. Introducció a les Xarxes Neuronals
5. Introducció al Aprenentatge Profund (Deep Learning, DL)
6. Xarxes Neuronals Convolucionals (CNN)
7. Entrenament: preprocessat de dades, inicialització, optimització del gradient
8. Classificació d'imatges
9. Entendre i visualització de CNNs
10. Mètodes eficients per a l'Aprenentatge Profund

Metodologia

Sesions supervisades:

- Sesions magistrals, on els professors explicaran continguts generals dels diferents temes. La majoria seran necessaris per a la resolució de problemes.

Sessions dirigides:

- Sesions de projecte, on es presentaran i discutiran els objectius i problemes dels projectes. Els estudiants hauran d'interactuar amb el coordinador de projecte sobre els problemes sorgits i les idees per a resoldre'ls. (Approx. 1 hora/setmana)
- Sesions de presentació, on els estudiants faran una presentació oral sobre como han solucionat el projecte i sobre els resultats obtinguts.
- Sesió d'exàmen, on els estudiants seran avaluats individualment, demostrant l'adquisició dels coneixements desenvolupats i la capacitat de resolució de problemes associats.

Treball autònom:

- Els estudiants hauran d'estudiar i treballar autònomament amb els materials derivats de les classes magistrals i de les sessions de projecte.
- Els estudiants treballaran en grup per a resoldre els problemes plantejats en els projectes amb els següents entregables:
 - Codi
 - Informe
 - Presentació oral

Activitats formatives

Títol	Hores	ECTS	Resultats d'aprenentatge
Tipus: Dirigides			
Sessions teòriques	20	0,8	3, 4, 8
Tipus: Supervisades			
Sessions de seguiment de projectes	8	0,32	1, 2, 3, 4, 5, 6, 7, 8, 9, 10
Tipus: Autònomes			
Treball autònom	113	4,52	1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Avaluació

La Nota final d'aquest mòdul es calcula amb la següent fórmula:

$$\text{Nota final} = 0.4 \times \text{Examen} + 0.55 \times \text{Projecte} + 0.05 \times \text{Assistència}$$

on,

Exàmen: és la nota obtinguda en l'examen final del mòdul (ha de ser ≥ 3)

Assistència: és la nota derivada de l'assistència a les sessions (mínim 70%)

Projecte: és la nota que posa el coordinador del projecte d'acord amb els resultats del seguiment setmanal del projecte i de les entregues. Tot d'acord amb criteris específics com ara:

- Participació en les sessions de discussió i en el treball en equip (avaluació entre-membres)
- Entrega de les parts obligatòries i opcionals del projecte.
- Desenvolupament del codi (estil, comentaris, etc.)
- Informe (justificació de les decisions preses en el desenvolupament del projecte)
- Presentació (Presentació i demostració del projecte desenvolupat)

Activitats d'avaluació

Títol	Pes	Hores	ECTS	Resultats d'aprenentatge
Assistència a sessions	0.05	0,5	0,02	1, 3, 4, 5, 8
Exàmen	0.4	2,5	0,1	1, 2, 4, 5, 7, 8, 10
Projecte	0.55	6	0,24	1, 2, 3, 4, 5, 6, 7, 8, 9, 10

Bibliografia

Articles en revistes:

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.

Llibres:

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 tècnics:

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