

## PROGRAMA DE CURSO

Código	Nombre			
MDS7203	Modelos Generativos Profundos			
Nombre en Inglés				
Deep Generative Models				
SCT	Unidades Docentes	Horas de Cátedra	Horas Docencia Auxiliar y Laboratorio	Horas de Trabajo Personal
6	10	1.5	1.5	7
Requisitos			Carácter del Curso	
MDS7104/EL4106/MA5204/MA4402/AUTOR			Electivo Magíster en Ciencia de Datos	
Resultados de Aprendizaje				
El objetivo del curso es presentar los fundamentos teóricos y algorítmicos que permiten construir y entrenar computacionalmente diferentes modelos generativos profundos, explorando técnicas contemporáneas de generación de datos en diversos dominios, y aplicarlos en la resolución de problemas prácticos. A lo largo del curso, los estudiantes adquirirán conocimientos sobre una variedad de modelos generativos, dominarán las hipótesis subyacentes de estos, comprenderán su ventajas y desventajas en distintos contextos, y aprenderán a evaluar la pertinencia de un modelo generativo (y cuál en específico) según el contexto. Finalmente, serán capaces de implementar y ajustar los modelos a través de laboratorios prácticos.				

Metodología Docente	Evaluación General
<p>La metodología consiste en:</p> <ul style="list-style-type: none"> <li>- Clases presenciales participativas, donde se expondrá la teoría detrás de cada técnica generativa junto con sus hipótesis, ventajas y desventajas conceptuales (1.5h/semana).</li> <li>- Trabajo dirigido, donde los estudiantes implementarán los modelos para comprender a cabalidad su pertinencia. Esto es en conjunto con los auxiliares (1.5h/semana).</li> <li>- Trabajo personal, donde los estudiantes revisan el material previo a la clase y completan el trabajo práctico después de ella (7h/semana).</li> </ul>	<p>La evaluación incluirá:</p> <ul style="list-style-type: none"> <li>- Estudio de los contenidos previo a la cátedra.</li> <li>- Asistencia y participación en las clases expositivas</li> <li>- Desempeño en los laboratorios prácticos a desarrollar en cada unidad luego de la clase expositiva.</li> <li>- Proyecto final del curso: aplicación práctica con su debida justificación teórica, inspirada en alguna técnica vista en el curso o alguna variante de conceptos estudiados en el curso.</li> <li>- Capacidad para presentar su estudio en el curso tanto de forma oral como escrita</li> </ul>



#### Unidades Temáticas

Número	Nombre de la Unidad	Duración en Semanas
1	Introducción	3
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> <li>- Definición de modelo generativo y ejemplos. Diferencia con modelo discriminativo.</li> <li>- Modelos gráficos, inferencia en grafos.</li> <li>- Principio y métodos básicos de <i>sampleo</i> de variables aleatorias:</li> </ul>	<p>Al final de la unidad se espera que el estudiante:</p> <ul style="list-style-type: none"> <li>- Entienda el principio de los modelos generativos y su diferencia con otros tipos de modelos en aprendizaje de máquinas.</li> </ul>	[1-4]

<p>métodos de Monte Carlo.</p> <ul style="list-style-type: none"> <li>- Distancias entre distribuciones de probabilidad y transporte óptimo.</li> </ul>	<ul style="list-style-type: none"> <li>- Identifique modelos generativos básicos y aquellos del estado del arte.</li> <li>- Entiendo la notación de en base a modelos gráficos.</li> <li>- Comprenda los fundamentos matemáticos necesarios para generar realizaciones desde una distribución de probabilidad.</li> <li>- Entienda y calcule distintas divergencias o distancias entre distribuciones de probabilidad en el contexto del aprendizaje de máquinas.</li> </ul>	
---	--	--

Número	Nombre de la Unidad	Duración en Semanas
2	Autoencoders Variacionales	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> <li>- Autoencoders deterministas y concepto de espacio latente. Relación con compresión.</li> <li>- Esquema de un autoencoder variacional</li> <li>- Inferencia variacional: Derivación de la cota inferior de la evidencia (ELBO).</li> <li>- Truco de la re-parametrización.</li> </ul>	<p>Al final de la unidad se espera que el estudiante:</p> <ul style="list-style-type: none"> <li>- Entienda la naturaleza de los modelos basados en autoencoders y el rol que cumple en estos un espacio latente de menor dimensionalidad.</li> <li>- Comprenda la teoría detrás del objetivo de optimización de un autoencoder variacional, así como también su proceso de <i>muestreo</i>.</li> <li>- Sea capaz de implementar un autoencoder variacional en un dataset de magnitud adecuada.</li> </ul>	[1], [5-7]

Número	Nombre de la Unidad	Duración en Semanas
3	Modelos basados en score	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía

<ul style="list-style-type: none"> <li>- Modelos basados en energía: Definición y pertinencia.</li> <li>- Definición y aproximación de la función de <i>score</i>.</li> <li>- <i>Sampleo</i> en base la función <i>score</i> y rol del ruido.</li> <li>- Generalización basada en ecuaciones derivadas estocásticas.</li> </ul>	<p>Al final de la unidad se espera que el estudiante:</p> <ul style="list-style-type: none"> <li>- Entienda el modelamiento de una distribución de probabilidad no-normalizada a través de una función de energía y su función <i>score</i>.</li> <li>- Comprenda la importancia de la inyección del ruido en la aproximación de la función <i>score</i> y su uso a través de ecuaciones derivadas estocásticas (EDEs).</li> <li>- Implemente un modelo generativo simple basado en <i>score</i>.</li> </ul>	[1],[8-10]
---	--	------------



Número	Nombre de la Unidad	Duración en Semanas
4	Modelos de difusión	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> <li>- Ruido <i>forward</i> y <i>backward</i>.</li> <li>- Concepto de autoencoder variacional jerárquico markoviano.</li> <li>- Eliminación de ruido a través de la cota inferior de la evidencia (ELBO).</li> <li>- Classifier guidance, classifier-free guidance para controlar el muestreo en un modelo de difusión, y su relación con el concepto de temperatura.</li> </ul>	<p>Al final de la unidad se espera que el estudiante:</p> <ul style="list-style-type: none"> <li>- Conozca las componentes de un entrenamiento basado en procesos de difusión y sus ventajas.</li> <li>- Identifique la equivalencia entre los modelos difusión y modelos de <i>score</i> basados en EDEs.</li> <li>- Implemente y modifique adecuadamente modelos de difusión para controlar la calidad y diversidad de la generación de datos.</li> </ul>	[1],[6] ,[11-15]

Número	Nombre de la Unidad	Duración en Semanas
5	Tópicos Complementarios	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> <li>- Componentes de una Red Generativa Antagónica (GAN).</li> <li>- Redes generadora y discriminadora</li> <li>- Formulación de las funciones de costo de ambas redes.</li> <li>- Divergencia de Jensen-Shannon.</li> <li>- Wasserstein GANs.</li> <li>- Modelos generativos basados en flujo.</li> <li>- Flujos normalizadores a través de transformaciones invertibles.</li> <li>- Sampleo y estimación de densidad en modelos de flujo.</li> </ul>	<p>Al final de la unidad se espera que el estudiante:</p> <ul style="list-style-type: none"> <li>- Conozca los componentes de una Red Generativa Antagónica (o GAN).</li> <li>- Comprenda la dinámica de entrenamiento competitiva entre las redes generadora y discriminadora en un modelo GAN.</li> <li>- Conozca variantes de redes generativas antagónicas basadas en principios del transporte óptimo.</li> <li>- Entienda el principio y relevancia de las transformaciones de una distribución de base para la expresividad de modelos de flujo.</li> <li>- Use cambio de variables para estimar la verosimilitud de datos según modelos de flujo.</li> </ul>	[1], [20], [23-26]

Número	Nombre de la Unidad	Duración en Semanas
6	Modelos Auto-regresivos	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
<ul style="list-style-type: none"> <li>- Formulación de un modelo auto-regresivo en base a probabilidades condicionales.</li> <li>- Módulo de atención neuronal.</li> <li>- Arquitectura Transformer solo decoder.</li> <li>- Decodificación de secuencias.</li> <li>- Modelos autor-regresivos para la generación de imágenes PixelCNN</li> </ul>	<p>Al final de la unidad se espera que el estudiante:</p> <ul style="list-style-type: none"> <li>- Entienda el principio matemático detrás de los modelos auto-regresivos y su aplicación al modelamiento generativo.</li> <li>- Conozca las arquitecturas neuronales usadas en modelos del estado del arte y sepa usarlas para desarrollar un generador de texto.</li> </ul>	[16-19] [21-22]

Número	Nombre de la Unidad	Duración en Semanas
7	Proyecto final	2
Contenidos	Resultados de Aprendizajes de la Unidad	Referencias a la Bibliografía
- Desarrollo autónomo de un modelo capaz de generar datos.	Al final de la unidad se espera que el estudiante: - Aplique un modelo reciente que genere datos usando técnicas aprendidas en el curso o derivadas de estas. - Explique de manera clara tanto sus justificaciones teóricas como resultados.	

## Referencias

- [1] Tomczak, J. M. (2021). *Deep Generative Modeling*. Cham: Springer International Publishing. [Online](#)
- [2] Murphy, K. An introduction to graphical models. Technical Report, University of British Columbia (2001). [Online](#)
- [3] Andrieu, C., de Freitas, N., Doucet, A. & Jordan, M. An Introduction to MCMC for Machine Learning. *Machine Learning* vol 50, p. 5–43 (2003) [Online](#)
- [4] Peyré, P. & Cuturi, M. *Computational Optimal Transport* (2018). [Online](#)
- [5] Kingma, D. P., & Welling, M. (2019). An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, 12(4), 307-392. [Online](#)
- [6] Luo, C. (2022). Understanding diffusion models: A unified perspective. arXiv preprint arXiv:2208.11970.
- [7] A. Amini (2023) - Deep Generative Models ([lecture notes](#))
- [8] Song, Y., & Ermon, S. (2019). Generative Modeling by Estimating Gradients of the Data Distribution. *Advances in Neural Information Processing Systems*, 32. [Online](#)
- [9] Song, Y., Sohl-Dickstein, J., Kingma, D. P., Kumar, A., Ermon, S., & Poole, B. (2021, January 12). Score-Based Generative Modeling through Stochastic Differential Equations. *International Conference on Learning Representations*. [Online](#)
- [10] Yang Song (2021) - Generative Modeling by Estimating Gradients of the Data Distribution. [Online](#)
- [11] Ho, J., Jain, A., & Abbeel, P. (2020). Denoising Diffusion Probabilistic Models. *Advances in Neural Information Processing Systems*, 33, 6840–6851. [Online](#)
- [12] Dhariwal, P., & Nichol, A. (2021). Diffusion Models Beat GANs on Image Synthesis. *Advances in Neural Information Processing Systems*, 34, 8780–8794. [Online](#)
- [13] Ho, J., & Salimans, T. (2021, December 8). Classifier-Free Diffusion Guidance. *NeurIPS 2021 Workshop on Deep Generative Models and Downstream Applications*. [Online](#)
- [14] Nichol, A. Q., Dhariwal, P., Ramesh, A., Shyam, P., Mishkin, P., McGrew, B., Sutskever, I., & Chen, M. (2022). GLIDE: Towards Photorealistic Image Generation and Editing with Text-Guided Diffusion Models. *Proceedings of the 39th International Conference on Machine Learning*, 16784–16804. [Online](#)
- [15] Sander Dieleman (2022) -Guidance: a cheat code for diffusion models ([Blog post](#))
- [16] Bahdanau, D., Cho, K. H., & Bengio, Y. (2015, January). Neural machine translation by jointly learning to align and translate. In *3rd International Conference on Learning Representations, ICLR 2015*. [Online](#)

- [17] Weng, L. (2018). Attention? Attention!. ([Blog post](#))
- [18] Vaswani, A. et al. Attention is all you need. *NeurIPS* 2017. [Online](#)
- [19] Phuong, M. & Hutter, M. Formal algorithms for transformers. DeepMind Technical Report (2022) [Online](#)
- [20] Alec Radford, Luke Metz, Soumith Chintala: Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks. ICLR (Poster) 2016. [Online](#)
- [21] Van den Oord, A., Kalchbrenner, N., Espeholt, L., Vinyals, O., & Graves, A. (2016). Conditional image generation with pixelcnn decoders. *Advances in neural information processing systems*, 29. [Online](#)
- [22] Salimans, T., Karpathy, A., Chen, X., & Kingma, D. P. (2017). Pixelcnn++: Improving the pixelcnn with discretized logistic mixture likelihood and other modifications. ICLR (Poster) 2017. [Online](#)
- [23] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2020). Generative adversarial networks. *Communications of the ACM*, 63(11), 139-144. [Online](#)
- [24] Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein generative adversarial networks. Proceedings of the 34th International Conference on Machine Learning - Volume 70, 214–223. [Online](#)
- [25] Rezende, D., & Mohamed, S. (2015, June). Variational inference with normalizing flows. In *International conference on machine learning* (pp. 1530-1538). PMLR. [Online](#)
- [26] Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., & Lakshminarayanan, B. (2021). Normalizing flows for probabilistic modeling and inference. *The Journal of Machine Learning Research*, 22(1), 2617-2680. [Online](#)



<b>Vigencia desde:</b>	Primavera 2023
<b>Elaborado por:</b>	Felipe Tobar
<b>Revisado por:</b>	Comité Académico MDS