

UTN Regional Paraná - Maestría en Minería de Datos

Plan de Tesis

Octubre 20223

Directores: Dr. Matías Gerard y Dr. Leandro Vignolo (CONICET-UNL)

Título: “Aplicación de autoencoders variacionales para mejorar los procesos de optimización evolutiva multiobjetivo”

Alumno: Lic. Claudio Sebastián Castillo (UTN)

Fundamentación y Justificación del tema

Los algoritmos evolutivos (en adelante AE) son métodos heurísticos de optimización inspirados en la evolución natural, diseñados para encontrar soluciones en espacios de búsqueda complejos. A diferencia de los métodos de optimización exhaustivos (e.g. *full search*), los AE son particularmente efectivos en espacios de búsqueda discretos, ruidosos o cuando la función objetivo es desconocida o no diferenciable (Goldberg, David E. 1989). Utilizando técnicas *evolutivas* estos algoritmos generan iterativamente soluciones a partir de una población de candidatos, de manera similar a cómo la evolución natural optimiza características biológicas a lo largo de nuevas generaciones. Aplicados a problemas multiobjetivo sus resultados regularmente implican soluciones cercanas al óptimo, que mantienen un buen compromiso en la satisfacción de múltiples requerimientos (Jiao et al. 2023).

La robustez de los AE está determinada, como bien sostiene Goldberg (1989), por una serie de características distintivas, que fortalecen su configuración de búsqueda, a saber: a) operan sobre un espacio *codificado* del problema y no sobre el espacio en su representación original; b) avanzan evaluando una *población de soluciones* y no soluciones individuales; c) tienen como guía una *función de aptitud* que no incluye derivación u otras funciones de cálculo; y d) suponen *reglas probabilísticas de transición* (operadores estocásticos) y no técnicas determinísticas. Estas características permiten a los AE superar restricciones que tienen otros métodos de optimización, condicionados -por ejemplo- a espacios de búsqueda continuos, diferenciables o unimodales. Por ello, su aplicación se ha extendido notablemente fuera del campo de la optimización, aplicándose en distintas tareas (Vie, Kleinnijenhuis, and Farmer 2021) y a lo largo de diversas industrias (Jiao et al. 2023).

A pesar de su robustez los AE se enfrentan a problemas desafiantes cuando se aplican a *datasets* de alta dimensionalidad y bajo número de muestras. En efecto, como sucede con gran parte de los algoritmos de aprendizaje automático (en adelante AA), en espacios de alta dimensionalidad la cardinalidad del conjunto de soluciones candidatas crece de manera exponencial, y los problemas se vuelven computacionalmente intratables. Cabe destacar que para un conjunto de n características es posible determinar un total de n^2 posibles soluciones, espacio que constituye un dominio de búsqueda difícil de cubrir aún con n conservadores.¹ Al mismo tiempo, la escasez de datos afecta gravemente el proceso de optimización, limitando la capacidad informativa de la función objetivo (Hastie, Tibshirani, and Friedman, Jerome 2009). En este sentido, la reconocida *maldición de la dimensionalidad*² también acecha a los AE pues a medida que las soluciones ganan complejidad (atributos) se requieren mayor cantidad

¹Por ejemplo para un dataset de 20 características (atributos) el número total de subconjuntos a evaluar supera el millón de posibles candidatos, específicamente: 1.048.576.

²Termino acuñado por Richard Bellman (1961) y que en el mundo del *aprendizaje automático* refiere a los problemas y desafíos que presentan los datos organizados en espacios de altas dimensiones (múltiples atributos). En tales espacios, la cantidad de datos necesarios para proporcionar una representación estadísticamente confiable crece exponencialmente pues toda representación tendrá gran dispersión entre sus datos. Esto significa que para cualquier conjunto de individuos la distancia promedio entre ellos aumenta a medida que aumentan las dimensiones, haciendo que cada uno se separe de los otros cada vez más.

de muestras (datos) para que la función objetivo aporte direcciones valiosas en el espacio de búsqueda y escape a soluciones subóptimas (Hamada and Goto 2018).

Por lo dicho, la disponibilidad de datos muestrales es un aspecto crítico de los algoritmos de AA en general y de los AE en particular, jugando un papel central en muchas situaciones prácticas. Así pues, resulta clave en contextos de optimización en el que se emplean arquitecturas neuronales profundas (Alzubaidi et al. 2023), pero también en otros escenarios donde, sin llegar a modelos muy complejos, se dispone de un conjunto de datos cuyas clases se encuentran desbalanceadas Blagus and Lusa (2013). La distribución inequitativa de datos entre clases es frecuente en sectores donde los eventos excepcionales son aquellos que despiertan mayor interés (e.g. sectores como salud, seguridad, telecomunicaciones y finanzas, son algunos ejemplos (Fajardo et al. 2021)). Mientras que las redes neuronales profundas y más ampliamente el tamaño de los modelos neuronales no han parado de crecer en aptitud y complejidad.³

Todo esto ha impulsado la búsqueda de estrategias de **aumentación** de datos que permitan superar los problemas derivados de su escasez y trabajar de manera intensiva con la información disponible (Gm et al. 2020). Entre esas estrategias el uso de Autoencoders Variacionales (en adelante AV) ocupa un lugar prominente. En efecto, los AV constituyen modelos generativos (redes neuronales profundas con arquitectura *encoder-decoder*) capaces de aprender una representación latente de datos disponibles y generar nuevas muestras con las mismas características fundamentales (i.e. similar distribución conjunta de probabilidad) que los datos originales (Kingma and Welling 2019). Esta capacidad generativa, fundada en la aptitud para modelar la distribución de probabilidad de ciertos datos, es particularmente efectiva por el hecho de que prescinde de fuertes supuestos estadísticos a los que adscriben otros modelos y por su *escalabilidad* gracias a emplear *retropropagación* como estrategia de optimización Kingma and Welling (2019). Hoy los AV son ampliamente utilizados en biología molecular, química, procesamiento de lenguaje natural, astronomía, entre otros (Ramchandran et al. 2022).

Por lo dicho hasta aquí, la posibilidad de expandir el conjunto de datos mediante el uso de AV, abre alternativas para mejorar el proceso de optimización de los AE. Esta contribución no solo estaría en permitir una mayor exploración del espacio de soluciones aplicando operadores evolutivos, sino también en mitigar riesgos de sobreajuste y convergencia prematura a soluciones subóptimas. De este modo, la integración de ambas tecnologías ofrece un enfoque prometedor para abordar escenarios de alta dimensionalidad y muestras escasas. A la fecha de publicación de la presente no hemos encontrado experiencias publicadas de aplicación general de AV en el ámbito de AE. En la medida que esto sea así creemos que nuestro aporte a la comunidad de investigadores y practicantes de la disciplina estará en proveer información y experimentación sobre la combinación de estos algoritmos largamente utilizados.

pendiente: introducir GA en EA (Vignolo and Gerard 2017)

³Baste citar como ejemplo muy popular en estos días al modelo *Transformer* y su enorme impacto dentro del mundo del AA y el NLP.

Estado del arte

AV en general

- (Fajardo et al. 2021)
- (Kwarciak and Wodzinski 2023)
- (Leelarathna et al. 2023)
- (Ai et al. 2023)
- (Tran, Bader, and Lundgren 2022)
- (Khmaissia and Frigui 2023)

AV en el contexto del AE

- (Martins, Rocha, and Pereira 2022)
- (Wu, Cao, and Qi 2023)

GA y búsqueda local

- (Vignolo and Gerard 2017)

Definición del problema

La escasez de datos muestrales en procesos de optimización evolutiva multiobjetivo y su posible solución a partir de técnicas de aumentación de datos mediante autoencoders variacionales.

Objetivos

General:

1. Investigar la eficacia de los autoencoders variacionales en la aumentación de datos para mejorar los procesos de optimización evolutiva multiobjetivo.

Específicos:

1. Implementar una arquitectura que integre autoencoders variacionales en la etapa inicial de un algoritmo evolutivo para la generación aumentada de población en vistas a la optimización.
2. Implementar una arquitectura que integre autoencoders variacionales en el contexto de optimización evolutiva, particularmente en la etapa de evolución de poblaciones, para favorecer el proceso de búsqueda de subconjuntos de características potencialmente valiosas.

Metodologías

Para cumplir los objetivos propuestos compararemos dos algoritmos genéticos evolutivos evaluando sus desempeños y resultados. Por un lado implementaremos un algoritmo genético *clásico* aplicado a conjuntos de datos conocidos dentro de la comunidad científica y en condiciones regulares de procesamiento para la resolución de problemas de clasificación. Por el otro, aplicaremos a los mismos conjuntos un variante *novedosa* de algoritmo genético, integrando un módulo de aumentación de datos a partir de la intervención de *autoencoders variacionales*. Para evaluar esta variante de algoritmo genético realizaremos diversos experimentos con ajustes en el diseño de la arquitectura que consistirán en la integración del módulo de aumentación de datos en dos puntos: 1) previo al inicio del proceso evolutivo y 2) durante el proceso evolutivo del algoritmo genético. Para la comparación de ambas implementaciones elegiremos como métricas la precisión (*accuracy*) en la clasificación y el tamaño final de características seleccionadas. A continuación detallamos los aspectos técnicos vinculados a los elementos mencionados, a saber: algoritmos, datos, arquitecturas y métricas para la evaluación.

1. **Algoritmos:** Para el presente trabajo usaremos algoritmos genéticos (AG) como método evolutivo⁴ debido a la posibilidad que brinda este algoritmo de emplear codificación binaria y permitir así una representación intuitiva del espacio de características (Vignolo and Gerard 2017). Para aumentación de datos utilizaremos *autoencoders variacionales* (AV) como instancia generativa (Kingma and Welling 2019).

1. Los AGs constituyen una de las herramientas más estudiadas e implementadas dentro de los métodos evolutivos Kramer (2017), dada su capacidad para encontrar soluciones en espacios de búsqueda complejos (Vignolo and Gerard 2017). El procedimiento de búsqueda de los AGs opera evolucionando una población de individuos que consisten en cromosomas que codifican el espacio de soluciones. Dicha evolución -al igual que la evolución natural- sucede a través de operadores (funciones) de selección, variación (mutación y cruce) y reemplazo que transforman el material genético disponible: los individuos más aptos sobreviven y se reproducen, mientras que los menos aptos desaparecen⁵. Esta aptitud -que imita la presión selectiva de un entorno natural- se evalúa mediante la aplicación de una función objetivo

⁴Otros métodos evolutivos robustos, como por ejemplo el *enjambre de partículas* (PSO) y *optimización de colonia de hormigas* (ACO), típicamente utilizan codificación basada en números reales por lo que constituyen opciones menos adecuadas al problema que enfrentaremos en este trabajo.

⁵Como su nombre lo indica el operador de selección determina la elegibilidad de un individuo para sobrevivir y reproducirse en función de su aptitud para resolver un problema. En el contexto de los AGs esta aptitud no es otra cosa que el puntaje que obtiene un individuo evaluado en una función objetivo. Por su parte los operadores de variación tienen como función combinar la información genética de individuos (cruce) y alterar aleatoriamente sus cromosomas (mutación), promoviendo transformaciones en el material genético global con sesgo hacia mejorar la aptitud poblacional para resolver un problema. La variación equivale a la búsqueda natural por mejorar las adaptaciones de los individuos a su entorno. Finalmente el operador de reemplazo mantiene la población constante, sustituyendo individuos poco aptos por aquellos de mayor aptitud. Estos operadores se combinan en ciclos iterativos que se repiten hasta satisfacer un criterio de terminación deseado (por ejemplo, un número predefinido de generaciones o un valor de aptitud) (Vignolo and Gerard 2017).

(específica del problema) a cada individuo a partir de la información decodificada de sus cromosomas. Este método heurístico de búsqueda tendrá en nuestro trabajo dos configuraciones: una *clásica* sin aumentación de datos y una *novedosa* con aumentación de datos aplicando *autoencoders variacionales* (AV).

2. Los AVs son modelos generativos (redes neuronales profundas con arquitectura *encoder-decoder*) capaces de aprender una representación latente de datos disponibles y generar nuevas muestras de similares características a los datos originales. Estos modelos se basan en el supuesto de que cualquier dato disponible, por ejemplo x , se genera mediante un proceso aleatorio que involucra una variable latente z . Bajo ese supuesto, el modelo procede tomando como muestra una observación x de la distribución previa de probabilidad $p(x)$, que luego se utiliza para tomar una observación z de la distribución condicional $p(z|x)$. El objetivo del modelo es obtener *estimaciones de máxima verosimilitud* (también llamadas *estimaciones de máxima probabilidad a posteriori*) de los parámetros en situaciones donde tanto la verosimilitud marginal $p(x) = \int p(x|z)p(z)dz$ como la probabilidad *a posteriori* $p(z|x)$ son intratables (Kingma and Welling 2019). Para eso, utiliza la distribución $p(z|x)$ como una aproximación al intratable $p(z)$, maximizando el *límite inferior variacional* para $p(z)$.

2. **Conjuntos de Datos.** Se llevarán a cabo experimentos utilizando cuatro conjuntos de datos con alta dimensionalidad y diferentes características, incluyendo un diverso número de características/atributos, ejemplos y clases/categorías.

1. *Madelon*: Este es un conjunto de datos artificial con 500 características, donde el objetivo es un XOR multidimensional con cinco características relevantes. Fue creado para el desafío de Selección de Características NIPS 2003⁶, y está disponible en el Repositorio UCI⁷. De las 495 características restantes, 15 corresponden a combinaciones lineales de las cinco relevantes, y las otras 480 son características de ruido. Madelon es un problema de clasificación de dos clases con variables de entrada binarias dispersas. Las dos clases están equilibradas, y los datos se dividen en conjuntos de entrenamiento y prueba.
2. *Leukemia*: El análisis de datos de expresión génica obtenidos de micro-datos de ADN se estudia en Golub (1999) para la clasificación de tipos de cáncer. Construyeron un conjunto de datos con 7129 mediciones de expresión génica en las clases ALL (leucemia linfocítica aguda) y AML (leucemia mielogénica aguda). El problema es distinguir entre estas dos variantes de leucemia (ALL y AML). Los datos se dividen originalmente en dos subconjuntos: un conjunto de entrenamiento y un conjunto de prueba independiente. El conjunto de entrenamiento consta de 38 muestras (27 ALL y 11 AML) de especímenes de médula ósea. El conjunto de prueba tiene 34 muestras (20 ALL y 14 AML), preparadas bajo diferentes condiciones experimentales e incluyendo 24 especímenes de médula ósea y 10 muestras de sangre.

⁶<http://clopinet.com/isabelle/Projects/NIPS2003/>

⁷<http://archive.ics.uci.edu/ml/datasets.html>

3. *GCM*: El conjunto de datos GCM fue compilado en Ramaswamy (2001) y contiene los perfiles de expresión de 198 muestras de tumores que representan 14 clases comunes de cáncer humano³. Aquí el enfoque estuvo en 190 muestras de tumores después de excluir 8 muestras de metástasis, y el preprocesamiento se realizó de acuerdo con [24]. Finalmente, cada matriz se estandarizó a una media de 0 y una varianza de 1 según [25]. El conjunto de datos consta de un total de 190 instancias, con 16063 atributos (biomarcadores) cada una, y distribuidos en 14 clases desequilibradas. En nuestros experimentos, los datos se separaron en tres conjuntos: entrenamiento (72 muestras), validación (72 muestras) y prueba (42 muestras). Con el objetivo de equilibrar las clases en el conjunto de entrenamiento y prevenir el sobreajuste del clasificador debido a la clase mayoritaria, las muestras se repitieron para obtener el mismo número para cada clase. Al final, se utilizaron un total de 168 muestras para el entrenamiento. Por el contrario, los conjuntos de validación y prueba están estratificados, manteniendo el desequilibrio de clases.
 4. *Gisette*: es un dataset creado para trabajar el problema de reconocimiento de dígitos escritos a mano (Isabelle Guyon 2004). Este conjunto de datos forma parte de los cinco conjuntos utilizados en el desafío de selección de características de NIPS 2003. Tiene 13500 observaciones y 5000 atributos. El desafío radica en diferenciar los dígitos ‘4’ y ‘9’, que suelen ser fácilmente confundibles entre sí. Los dígitos han sido normalizados en tamaño y centrados en una imagen fija de 28x28 píxeles. Para el desafío de selección de características, se modificaron los datos originales. Específicamente, se muestrearon píxeles al azar en la parte superior central de la característica que contiene la información necesaria para diferenciar el 4 del 9. Además, se crearon características de orden superior como productos de estos píxeles para sumergir el problema en un espacio de características de mayor dimensión. También se añadieron características distractoras denominadas “sondas”, que no tienen poder predictivo. El orden de las características y patrones fue aleatorizado.
3. **Arquitecturas**: Para nuestro trabajo utilizaremos dos arquitecturas: una *clásica* sin aumentación de datos y una *novedosa* con aumentación de datos aplicando *autoencoders variacionales* (AV). A su vez, dentro de esta segunda configuración ensayaremos ajustes en el diseño que consistirán en la integración del módulo de aumentación de datos en dos puntos: 1) previo al inicio del proceso evolutivo y 2) durante el proceso evolutivo del algoritmo genético. De esta forma buscamos evaluar el aporte generativo del modelo AV en distintas etapas del AG.
 4. **Métricas de evaluación**: Para la comparación de ambas implementaciones elegiremos como métricas la precisión en la clasificación (*accuracy*) y el tamaño final de características seleccionadas por el modelo genético. Generalmente los AGs constituye modelos *multiobjetivo* donde no solo interesa optimizar una métrica particular vinculada a su eficacia -por ejemplo precisión en su clasificación de individuos- sino también interesa satisfacer requerimientos de eficiencia -como por ejemplo minimizar el conjunto de características con que opera el modelo- (Jiao et al. 2023). Por esta razón nuestra evaluación tomará en cuenta ambas dimensiones en el desempeño de las soluciones propuestas.

Cronograma de trabajo

[pendiente]

Condiciones para el desarrollo

[pendiente]

Bibliográfia

- Ai, Qingzhong, Pengyun Wang, Lirong He, Liangjian Wen, Lujia Pan, and Zenglin Xu. 2023. “Generative Oversampling for Imbalanced Data via Majority-Guided VAE.” February 14, 2023. <http://arxiv.org/abs/2302.10910>.
- Alzubaidi, Laith, Jinshuai Bai, Aiman Al-Sabaawi, Jose Santamaría, A. S. Albahri, Bashar Sami Nayyef Al-dabbagh, Mohammed A. Fadhel, et al. 2023. “A Survey on Deep Learning Tools Dealing with Data Scarcity: Definitions, Challenges, Solutions, Tips, and Applications.” *Journal of Big Data* 10 (1): 46. <https://doi.org/10.1186/s40537-023-00727-2>.
- Blagus, Rok, and Lara Lusa. 2013. “SMOTE for High-Dimensional Class-Imbalanced Data.” *BMC Bioinformatics* 14 (1): 106. <https://doi.org/10.1186/1471-2105-14-106>.
- Doersch, Carl. 2021. “Tutorial on Variational Autoencoders.” January 3, 2021. <http://arxiv.org/abs/1606.05908>.
- Fajardo, Val Andrei, David Findlay, Charu Jaiswal, Xinshang Yin, Roshanak Housmanfar, Honglei Xie, Jiayi Liang, Xichen She, and D. B. Emerson. 2021. “On Oversampling Imbalanced Data with Deep Conditional Generative Models.” *Expert Systems with Applications* 169 (May): 114463. <https://doi.org/10.1016/j.eswa.2020.114463>.
- Gm, Harshvardhan, Mahendra Kumar Gourisaria, Manjusha Pandey, and Siddharth Swarup Rautaray. 2020. “A Comprehensive Survey and Analysis of Generative Models in Machine Learning.” *Computer Science Review* 38 (November): 100285. <https://doi.org/10.1016/j.cosrev.2020.100285>.
- Goldberg, David E. 1989. *Genetic Algorithms in Search, Optimization, and Machine Learning*. New York, NY, USA: Addison-Wesley.
- Golub, T. R., D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, et al. 1999. “Molecular Classification of Cancer: Class Discovery and Class Prediction by Gene Expression Monitoring.” *Science* 286 (5439): 531–37. <https://doi.org/10.1126/science.286.5439.531>.
- Hamada, Naoki, and Keisuke Goto. 2018. “Data-Driven Analysis of Pareto Set Topology.” In *Proceedings of the Genetic and Evolutionary Computation Conference*, 657–64. <https://doi.org/10.1145/3205455.3205613>.
- Hastie, Trevor, Robert Tibshirani, and Friedman, Jerome. 2009. *The Element of Statistical Learning: : Data Mining, Inference, and Prediction*. Second Edition. Springer.
- Isabelle Guyon, Steve Gunn. 2004. “Gisette.” UCI Machine Learning Repository. <https://doi.org/10.24432/C5HP5B>.
- Jiao, Ruwang, Bach Hoai Nguyen, Bing Xue, and Mengjie Zhang. 2023. “A Survey on Evolutionary Multiobjective Feature Selection in Classification: Approaches, Applications, and Challenges.” *IEEE Transactions on Evolutionary Computation*, 1–1. <https://doi.org/10.1109/TEVC.2023.3292527>.
- Khmaissia, Fadoua, and Hichem Frigui. 2023. “Confidence-Guided Data Augmentation for Improved Semi-Supervised Training.” February 21, 2023. <http://arxiv.org/abs/2209.08174>.
- Kingma, Diederik P., and Max Welling. 2019. “An Introduction to Variational Autoencoders.” *Foundations and Trends® in Machine Learning* 12 (4): 307–92. <https://doi.org/10.1561/>

2200000056.

- Kramer, Oliver. 2017. *Genetic Algorithm Essentials*. Vol. 679. Studies in Computational Intelligence. Cham: Springer International Publishing. <https://doi.org/10.1007/978-3-319-52156-5>.
- Kwarciak, Kamil, and Marek Wodzinski. 2023. “Deep Generative Networks for Heterogeneous Augmentation of Cranial Defects.” August 9, 2023. <http://arxiv.org/abs/2308.04883>.
- Leelarathna, Navindu, Andrei Margeloiu, Mateja Jamnik, and Nikola Simidjievski. 2023. “Enhancing Representation Learning on High-Dimensional, Small-Size Tabular Data: A Divide and Conquer Method with Ensembled VAEs.” June 27, 2023. <http://arxiv.org/abs/2306.15661>.
- Martins, Miguel, Miguel Rocha, and Vítor Pereira. 2022. “Variational Autoencoders and Evolutionary Algorithms for Targeted Novel Enzyme Design.” In *2022 IEEE Congress on Evolutionary Computation (CEC)*, 1–8. <https://doi.org/10.1109/CEC55065.2022.9870421>.
- Ramaswamy, Sridhar, Pablo Tamayo, Ryan Rifkin, Sayan Mukherjee, Chen-Hsiang Yeang, Michael Angelo, Christine Ladd, et al. 2001. “Multiclass Cancer Diagnosis Using Tumor Gene Expression Signatures.” *Proceedings of the National Academy of Sciences* 98 (26): 15149–54. <https://doi.org/10.1073/pnas.211566398>.
- Ramchandran, Siddharth, Gleb Tikhonov, Otto Lönnroth, Pekka Tiikkainen, and Harri Lähdesmäki. 2022. “Learning Conditional Variational Autoencoders with Missing Covariates.” March 2, 2022. <http://arxiv.org/abs/2203.01218>.
- Tran, Thanh, Sebastian Bader, and Jan Lundgren. 2022. “An Artificial Neural Network-Based System for Detecting Machine Failures Using Tiny Sound Data: A Case Study.” September 23, 2022. <http://arxiv.org/abs/2209.11527>.
- Vie, Aymeric, Alissa M. Kleinnijenhuis, and Doyne J. Farmer. 2021. “Qualities, Challenges and Future of Genetic Algorithms: A Literature Review.” September 13, 2021. <http://arxiv.org/abs/2011.05277>.
- Vignolo, Leandro D., and Matias F. Gerard. 2017. “Evolutionary Local Improvement on Genetic Algorithms for Feature Selection.” In *2017 XLIII Latin American Computer Conference (CLEI)*, 1–8. Cordoba: IEEE. <https://doi.org/10.1109/CLEI.2017.8226467>.
- Wong, Sebastien C., Adam Gatt, Victor Stamatescu, and Mark D. McDonnell. 2016. “Understanding Data Augmentation for Classification: When to Warp?” In *2016 International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, 1–6. Gold Coast, Australia: IEEE. <https://doi.org/10.1109/DICTA.2016.7797091>.
- Wu, Zhangkai, Longbing Cao, and Lei Qi. 2023. “eVAE: Evolutionary Variational Autoencoder.” January 1, 2023. <https://doi.org/10.48550/arXiv.2301.00011>.