EAUC: Evaluación Sensible al Sesgo en Regresión Diádica

Jorge Paz-Ruza, Amparo Alonso-Betanzos, Bertha Guijarro Berdiñas, Brais Cancela, and Carlos Eiras-Franco

Universidade da Coruña, Grupo LIDIA - CITIC
{j.ruza, amparo.alonso.betanzos, berta.guijarro, brais.cancela, carlos.eiras.franco}@udc.es

Resumen La regresión diádica, donde se predicen valores reales para pares de entidades, es fundamental en dominios como los sistemas de recomendación, las finanzas o la farmacología de precisión. En este trabajo resumimos nuestra investigación reciente sobre un sesgo sistemático presente en los modelos de regresión diádica, al que denominamos *eccentricity bias*. Este fenómeno desplaza las predicciones hacia los valores medios observados históricamente para usuarios e ítems, reduciendo significativamente la precisión en ejemplos excéntricos y derivando en sesgos sistemáticos al manejar de forma injusta estos ejemplos. Mostramos que métricas globales clásicas como RMSE o MAE no tienen la capacidad de detectar este comportamiento y presentamos EAUC (*Eccentricity-Area Under the Curve*), una nueva métrica orientada a evaluar dicho sesgo. Validamos experimentalmente EAUC en distintos dominios críticos y discutimos su utilidad para el desarrollo de modelos más justos.

Keywords: Regresión diádica · Fairness · Sistemas de recomendación · Evaluación · Sesgo algorítmico

1. Introducción

Los sistemas basados en datos diádicos modelan interacciones entre pares de entidades, típicamente usuarios e ítems. Este paradigma aparece en sistemas de recomendación, predicción de préstamos financieros, comercio internacional o farmacología personalizada. Formalmente, dado un usuario u y un ítem i, el objetivo consiste en predecir un valor real $\hat{r}_{ui}$ asociado a dicha interacción.

En numerosos escenarios, estas predicciones afectan directamente a personas y procesos sensibles, lo que convierte la equidad y ausencia de sesgos en un requisito fundamental. Regulaciones recientes como el AI Act europeo enfatizan precisamente la necesidad de evitar patrones discriminatorios en sistemas de IA.

En nuestro trabajo original [1], condensado en la Figura 1, identificamos un fenómeno sistemático en modelos de regresión diádica: las predicciones tienden a desplazarse hacia la media histórica observada de usuarios e ítems. Como consecuencia, los modelos obtienen buenos resultados en métricas globales como RMSE, pero fallan de forma severa en ejemplos excéntricos, precisamente

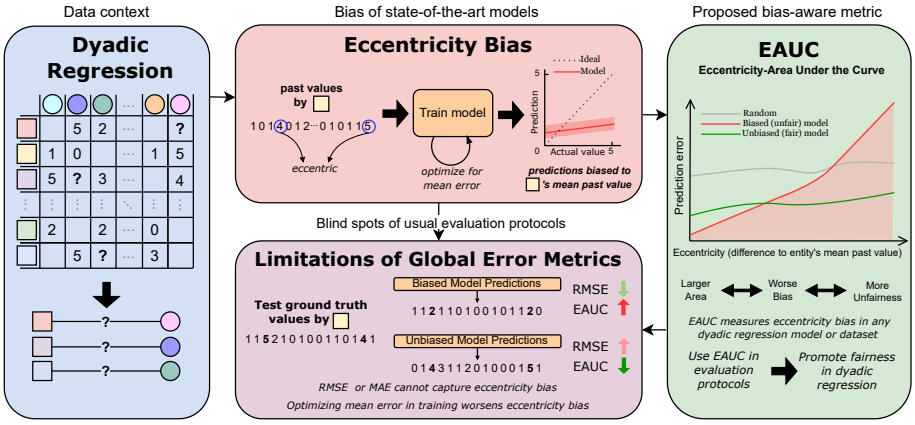


Figura 1: Resumen conceptual del sesgo de excentricidad y del uso de EAUC.

aquellos más relevantes desde la perspectiva de justicia algorítmica y robustez; denominamos a este comportamiento *sesgo de excentricidad* (*eccentricity bias*).

En el trabajo demostramos que las métricas tradicionales utilizadas históricamente en estos problemas son insuficientes para detectar este comportamiento, y en consecuencia proponemos EAUC (*Eccentricity-Area Under the Curve*), una nueva métrica diseñada específicamente para cuantificar el sesgo de excentricidad. Este artículo resume las principales contribuciones del trabajo original publicado previamente, enfatizando los aspectos conceptuales, metodológicos y experimentales más relevantes.

2. Sesgo de Excentricidad en Regresión Diádica

Sea $\mathcal{U}$ el conjunto de usuarios e $\mathcal{I}$ el conjunto de ítems. Un problema de regresión diádica (Fig. 1, izquierda) busca aprender una función $f : \mathcal{U} \times \mathcal{I} \to \mathbb{R}$ que prediga un valor real r_{ui} que caracterice la relación (u, i) (Fig. 1 superior).

La mayoría de modelos actuales optimizan pérdidas globales como MSE o RMSE, adecuadas para una regresión clásica. Sin embargo, en datos diádicos cada usuario e ítem posee distribuciones locales de valores observados diversas y no uniformes; en consecuencia, los modelos reciben una fuerte recompensa por aproximar simplemente el punto medio de las medias históricas de cada entidad.

Definimos el Valor Medio Diádico (DMV) como:

$$DMV_{ui} = \frac{\bar{r}_u + \bar{r}_i}{2}$$

donde $\bar{r}_u$ y $\bar{r}_i$ representan los valores medios históricos observados para el usuario y el ítem respectivamente. Experimentalmente observamos que todos los modelos considerados generan predicciones $\hat{r}_{ui}$ excesivamente próximas a DMV_{ui} incluso cuando el valor real r_{ui} se aleja significativamente de dicha media. Esto provoca errores especialmente altos en ejemplos excéntricos (Fig. 1 inferior).

3. EAUC: Evaluación Sensible al Sesgo en Regresión Diádica

Para detectar y cuantificar el *eccentricity bias* proponemos EAUC (*Eccentricity-Area Under the Curve*). La motivación principal de EAUC es que métricas globales como RMSE o MAE únicamente reflejan el error promedio total del modelo: un modelo puede obtener buenos resultados globales mientras falla sistemáticamente en ejemplos alejados de los respectivos DMV.

La excentricidad de un ejemplo (u, i) se define como:

$$Ecc_{ui} = |r_{ui} - DMV_{ui}|$$

donde r_{ui} es el valor real observado, y $\bar{r}_u$, $\bar{r}_i$ las medias históricas observadas del usuario u e item i. Intuitivamente, Ecc_{ui} mide cuánto se desvía un ejemplo respecto al comportamiento esperado según el historial de ambas entidades.

A continuación se calcula el error absoluto de predicción como $\epsilon_{ui} = |\hat{r}_{ui} - r_{ui}|$. Una vez obtenidos los pares $(Ecc_{ui}, \epsilon_{ui})$ para todos los ejemplos de test:

1. Los ejemplos se ordenan por excentricidad creciente.
2. Se construye la curva excentricidad-error.
3. Se calcula el área bajo dicha curva.
4. El resultado se normaliza al rango $[0, 1]$.

Un modelo ideal, capaz de predecir correctamente tanto ejemplos normales como excéntricos, obtiene $EAUC \approx 0$. Por el contrario, valores altos indican que el error crece rápidamente conforme aumenta la excentricidad, reflejando un fuerte sesgo hacia las medias históricas.

EAUC puede utilizarse de dos formas complementarias, ilustradas en la Figura 2: como una dimensión adicional de evaluación (Subfig. 2a) , o mediante el análisis de la curva excentricidad-error (Subfig. 2b).

4. Evaluación Experimental

En nuestro trabajo original, experimentamos con diez conjuntos de datos reales de regresión diádica pertenecientes a tres dominios: recomendación de productos, farmacología de precisión, y predicción de datos macroeconómicos y de finanzas. Evaluamos distintos modelos representativos del estado del arte, incluyendo Matrix Factorization (MF), Bayesian SVD++, GC-MC y Glocal-K. Como ejemplo, la Figura 3 muestra la relación observada entre excentricidad y error de predicción en un contexto de recomendación de productos.

Nuestros resultados muestran consistentemente que:

1. Todos los modelos estudiados presentan sesgo de excentricidad.
2. Un menor RMSE no implica necesariamente menor sesgo.
3. Algunos modelos obtienen incluso peor comportamiento que predictores aleatorios en ejemplos altamente excéntricos.
4. EAUC detecta adecuadamente este sesgo en todos los dominios analizados.

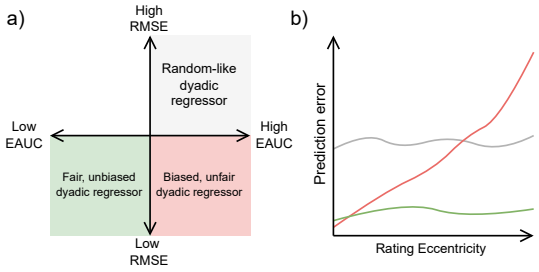


Figura 2: Interpretación y uso de EAUC.

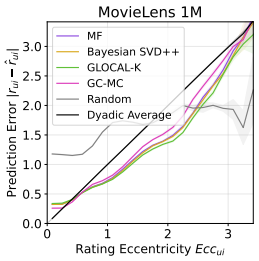


Figura 3: Relación entre excentricidad y error de predicción en el conjunto ML-1M.

5. Implicaciones y Conclusiones

Los resultados experimentales muestran que el *eccentricity bias* aparece de forma consistente en modelos modernos de regresión diádica y en dominios con fuerte impacto social, incluyendo recomendación, farmacología y finanzas. Aunque muchos modelos alcanzan buenos valores de RMSE o MAE, pueden presentar errores muy elevados en ejemplos excéntricos, precisamente aquellos más sensibles desde la perspectiva de fairness y robustez.

EAUC permite detectar este comportamiento analizando cómo evoluciona el error de predicción conforme los ejemplos se alejan de las medias históricas de usuarios e ítems. De este modo, complementa las métricas globales tradicionales incorporando una dimensión explícita de sesgo y capacidad de generalización.

Además de su utilidad evaluativa, EAUC podría guiar el desarrollo de modelos más justos mediante correcciones *post-hoc* o como termino a optimizar en entrenamiento. Adicionalmente, EAUC podría contribuir a comprender mejor los sesgos individuales o de grupo clásicos (e.g. aquellos demográficos o de popularidad), que podrían estar causados por sesgos de excentricidad y no de representación.

Declaración de intereses Los autores declaran no tener intereses en conflicto.

Agradecimientos Este trabajo está financiado por MICIU/AEI FEDER Una forma de hacer Europa (PID2023-147404OB-I00), Ministerio de Transformación Digital y Función Pública y 'Next-GenerationEU' /PRTR (TSI-100925-2023-1), y Xunta de Galicia (ED431C 2026/38, ED431G 2023/01).

Referencias

1. Paz-Ruza, J., Alonso-Betanzos, A., Guijarro-Berdiñas, B., Cancela, B., Eiras-Franco, C.: Beyond rmse and mae: Introducing eauc to unmask hidden bias and unfairness in dyadic regression models. IEEE Transactions on Neural Networks and Learning Systems **36**(11), 19619–19630 (2025). https://doi.org/10.1109/TNNLS.2025.3593059