

Systematic mapping of Predicting Weather Forecast Uncertainty with Machine Learning

Carmen Calvo-Olivera¹

¹Supercomputación Castilla y León (SCAYLE). Campus de Vegazana S/N, 24071 León, Spain

Abstract. *Dado el carácter incierto asociado a las predicciones meteorológicas, resulta necesario proporcionar un índice de verosimilitud para establecer si podemos considerar valiosa o no dicha predicción. El objetivo de este artículo es realizar una revisión sistemática de la bibliografía para establecer qué algoritmos de clasificación y conjuntos de datos ofrecen mejores resultados a la hora de determinar el índice de verosimilitud de modelos de predicción como el WRF (Weather Research and Forecasting).*

Keywords: *Weather Forecast, Uncertainty, Machine Learning, Systematic Mapping*

1. Introducción

Las predicciones meteorológicas son inherentemente inseguras. Por lo que, para muchas aplicaciones las previsiones sólo se consideran valiosas si existe una incertidumbre asociada a dicha predicción [Scher and Messori 2018]. Actualmente, el mejor método para proporcionar una estimación de confianza para las predicciones individuales es producir un conjunto de simulaciones numéricas del clima, lo cual resulta computacionalmente muy caro y requiere, por tanto, del uso de superordenadores.

[...]

Sin embargo, la literatura actual no proporciona suficiente información sobre la estimación de la incertidumbre asociada a una predicción, la cual puede suponer una importante vía de estudio en lo que respecta a la afianzamiento de la verosimilitud de los resultados de una predicción, por lo que es útil realizar revisiones de la literatura que aborden la estimación del índice de verosimilitud asociado a una predicción meteorológica. Esta razón es la principal motivación de este trabajo.

El objetivo del estudio que se ha realizado es proporcionar una visión objetiva del contexto tecnológico actual, en relación con la definición de la incertidumbre asociada a una predicción. Por lo tanto, el objetivo es responder a la siguiente pregunta de investigación:

RQ1 *¿Cuáles son las tendencias actuales, respecto al uso de algoritmos y conjuntos de datos, en términos de determinación de la incertidumbre asociada a una predicción meteorológica?*

Este trabajo está organizado de la siguiente manera: en la sección 2 se especifica el proceso seguido para la recopilación y filtrado del conjunto de artículos de estudio sobre el que se realiza la investigación; en la sección 3 se muestran los resultados obtenidos

en la búsqueda y construcción del conjunto de artículos de estudio, así como el análisis de los resultados obtenidos tras extraer la información de las diferentes muestras; por último, en la sección 4 se detallan las conclusiones y posibles aplicaciones futuras de la labor realizada en este estudio.

2. Metodología

Para llevar a cabo el estudio, se han seguido las recomendaciones de Kitchenham [Kitchenham et al. 2015], así como la guía de PRISMA [Moher et al. 2009]. La metodología se divide en tres fases: planificación de la búsqueda, proceso de búsqueda y selección de muestras, y extracción de datos y preparación del informe.

2.1. Plan de búsqueda

Una vez comprobado que no se ha realizado un estudio centrado en el campo de interés y que ha respondido a la pregunta de investigación planteada en este artículo, se ha planificado la búsqueda de los artículos que componen el conjunto de estudios con los que hemos trabajado. Se consideraron varias fuentes de información: Web of Science, IEEE Digital Library, Scopus.

2.2. Proceso de búsqueda

Una vez decididas las bases de datos en las que se realizarán las búsquedas, es necesario construir cadenas de búsqueda adecuadas que permitan obtener resultados satisfactorios. Para la realización de la búsqueda se han construido dos cadenas que se han aplicado a cada una de las bases de datos indicadas anteriormente. Las cadenas de búsqueda se han construido a partir de las palabras clave extraídas tras aplicar la estrategia PICOC [Roehrs et al. 2017, Schardt et al. 2007] a la pregunta de investigación planteada. Así pues, se construyeron dos cadenas de búsqueda:

SS1 (('weather forecast*') AND ('uncertainty') AND ('predict*') AND ('machine learning' OR 'classification' OR 'system') AND ('dataset'))

SS2 ('weather forecast*' OR 'rain forecast*' OR 'precipitation forecast*') AND ('predict*') AND ('uncertainty' OR 'likelihood index') AND ('machine learning' OR 'classification' OR 'training')

Utilizando las cadenas de búsqueda anteriores, se intentó reunir un conjunto de artículos que podría ser analizados en este documento. Otros de los filtros aplicados durante la búsqueda es el relacionado con la fecha de publicación del artículo dado que se trabaja en un campo relativamente moderno y que evoluciona rápidamente, como es la aplicación de técnicas de aprendizaje automático en la meteorología, debemos centrarnos sólo en artículos recientes. De esta manera, la búsqueda sólo se ha realizado para los artículos publicado entre 2017 y 2020. Por otro lado, sólo aquellos artículos cuyo acceso es libre han sido finalmente seleccionados.

Resultados para SS1:

- Web of Science: 63. De open access: 34
- IEEE: 11. De open access: 0
- SCOPUS: 11. De open access: 1

Resultados para SS2:

- Web of Science: 199. De open access: 95
- IEEE: 27. De open access: 1
- SCOPUS: 99. De open access: 31

2.3. Selección de resultados

Para seleccionar las muestras obtenidas, se planificó un proceso de filtrado. En primer lugar, se eliminaron los artículos duplicados, eliminando seis de los cuarenta y cuatro artículos iniciales. A continuación, se realizó un filtrado basado en la lectura del resumen de cada uno de los artículos restantes. Para decidir si un artículo en particular se marca como aceptado o rechazadas, se han establecido una serie de criterios de inclusión y exclusión. En el caso de que un artículo comparta un criterio de exclusión y un criterio de inclusión, se dará más peso al criterio de exclusión. De esta manera, se trata de eliminando aquellos artículos que cumplen al menos un criterio de exclusión o que no cumplen cualquiera de los criterios de inclusión.

Los criterios de inclusión establecidos, son los siguientes:

- CI1** Uso de datos meteorológicos. (?)
- CI2** El artículo trata cálculo de la incertidumbre asociada a las precipitaciones meteorológicas, ya sea centrándose en su estudio o simplemente haciendo referencia a él.
- CI3** El artículo utiliza técnicas de ML aplicadas al análisis de conjuntos de datos meteorológicos.
- CI4** El artículo utiliza técnicas de ML para construir un sistema que defina el índice de verosimilitud de una predicción.

En cuanto a los criterios de exclusión tenidos en cuenta:

- CE1** El artículo no emplea técnicas de ML.
- CE2** El artículo no pertenece al campo de la meteorología.
- CE3** El artículo en una revisión de la literatura.
- CE4** El artículo no tiene acceso gratuito.
- CE5** El artículo no está en inglés ni español.

Para lograr el objetivo de analizar aproximadamente diez artículos, se ha elaborado un cuestionario de evaluación de la calidad del artículo, basado en ciertos requisitos que son de interés en esta investigación. Este cuestionario se ha completado en base a la información obtenida después de la lectura completa de los artículos que han sido seleccionados en la fase anterior. El cuestionario que permitirá evaluar los artículos antes de pasar a la extracción de se compone de las siguientes preguntas:

- P1** ¿Menciona varios algoritmos de ML?
- P2** ¿Se calcula la incertidumbre asociada a una predicción?
- P3** ¿Aplica técnicas de ML?
- P4** ¿Especifica con qué dataset(s) trabaja?
- P5** ¿Adjunta métricas que validen los resultados obtenidos?

P6 ¿Especifica sobre qué fenómeno meteorológico se aplican los cálculos? - con esto quiero saber si lo aplica a la rain, a viento, a la temperatura -

Una vez seleccionadas las preguntas, es necesario establecer sus posibles respuestas y qué valor tendrá cada una de ellas. Para la construcción de este cuestionario, se ha intentado formular todas las preguntas para que puedan tener una respuesta del tipo "Sí", o "No", puntuando con 1 y 0 puntos, respectivamente. De este modo, la máxima puntuación que puede alcanzar un artículo es de seis puntos. Con este se ha establecido un umbral en cuatro puntos, para así incluir únicamente los artículos con una puntuación de 5 o 6, con el objetivo de que los artículos sean lo más completos posible. Dado que los artículos que no cumplen con P3 tampoco cumplirán con P1, esto implicará que esos artículos no se seleccionan automáticamente cuando no cumplen con la se indica la puntuación mínima.

2.4. Extracción de información

3. Resultados

4. Conclusiones

Referencias

- Kitchenham, B. A., Budgen, D., and Brereton, P. (2015). *Evidence-based software engineering and systematic reviews*, volume 4. CRC press.
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., and Group, P. (2009). Reprint—preferred reporting items for systematic reviews and meta-analyses: the prisma statement. *Physical therapy*, 89(9):873–880.
- Roehrs, A., Da Costa, C. A., da Rosa Righi, R., and De Oliveira, K. S. F. (2017). Personal health records: a systematic literature review. *Journal of medical Internet research*, 19(1):e13.
- Schardt, C., Adams, M. B., Owens, T., Keitz, S., and Fontelo, P. (2007). Utilization of the pico framework to improve searching pubmed for clinical questions. *BMC medical informatics and decision making*, 7(1):16.
- Scher, S. and Messori, G. (2018). Predicting weather forecast uncertainty with machine learning. *Quarterly Journal of the Royal Meteorological Society*, 144(717):2830–2841.