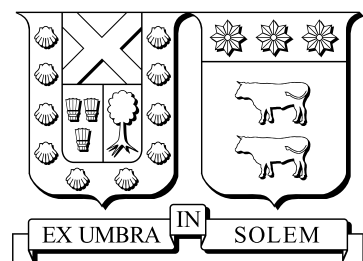


UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

Departamento de Informática

Valparaíso - Chile



**UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA**

**Representaciones homogéneas, heterogéneas y mixtas para
la estimación de potencial fotovoltaico urbano en Santiago**

Tesis sometida en cumplimiento parcial de los requisitos para el grado de
Magíster en Ciencias de la Ingeniería Informática de la Universidad Técnica
Federico Santa María

Magíster en Ciencias de la Ingeniería Informática

Alvaro Valderrama

Profesores guía

Prof. Dr. rer. nat. Héctor Allende Olivares (Tutor, UTFSM)

Prof. Dr. Carlos Valle Vidal (Co-tutor, UTFSM)

Julio 2023

UNIVERSIDAD TECNICA FEDERICO SANTA MARIA

Departamento de Informática
Valparaíso - Chile

Título

Representaciones homogéneas, heterogéneas y mixtas para la estimación de potencial fotovoltaico urbano en Santiago

Autor

Alvaro Valderrama

Tesis sometida en cumplimiento parcial de los requisitos para el grado de Magíster en Ciencias de la Ingeniería Informática de la Universidad Técnica Federico Santa María.

Dr. Héctor Allende O.
(Tutor)

Dr. Carlos Valle V.
(Co-tutor)

Dr. Mauricio Solar F.
(Evaluador Interno)

Dr. Rodrigo Palma B.
(Evaluador Externo)

Julio 2023

Agradecimientos

Primeramente, me gustaría agradecer el apoyo y confianza otorgado por mi profesor tutor durante este proceso, el que me ha permitido explorar esta área del conocimiento y, espero, responder a esa confianza. Igualmente, al profesor Carlos Valle, quien, luego de guiar mi proyecto de pregrado, me recomendó al profesor Allende, y luego siempre ha seguido empujando mis iniciativas, y ha co-guiado este trabajo con interés y compromiso. Y más generalmente, al grupo de investigación INCA, por la paciencia con la que escucharon mis ideas y la honestidad con la que aportaron sus opiniones y recomendaciones.

A mi familia, mi novio y mis amigos, agradecerles la compañía y apoyo estos años. A veces el mejor camino para solucionar un problema es dejar de pensar en el por unos instantes. Fueron más indispensables de lo que podría parecer.

Finalmente, a mi socio Marcelo. La idea de este proyecto surgió trabajando en una de nuestras muchas locuras. Incluso cuando de lejos algunas parecían imposibles, siempre me ayudó a ver caminos inesperados y soluciones divertidas, y, ante todo, a perseverar.

Índice general

1. Introducción	1
2. Contexto	3
3. Estado del Arte	5
4. El caso chileno	16
5. Problema de aprendizaje general	18
6. Convoluciones 1D	20
7. Propuesta e hipótesis	22
8. Experimentos	26
8.1. Conjunto de datos	26
8.2. Ajuste de hiperparámetros	26
8.3. Comparación redes optimizadas	28
8.4. Comparación costos computacionales	29
8.5. Representaciones internas	30
9. Contribuciones	32
10. Conclusiones	33
11. Trabajo futuro	35

Índice de figuras

- 7.1. Representación esquemática de la operación convolucional de una sola neurona sobre la representación heterogénea de una manzana particular. La definición matemática de la convolución presentada se encuentra descrita en la ecuación (7.5). Obtenida de Valderrama *et al.* [76], en inglés del original. 24

- 8.1. Presenta una representación esquemática de las tres arquitecturas sobre las que se realizó optimización de los hiperparámetros. Cada casilla comenzando con una C corresponde a una convolución, BN corresponde a una capa de *Batch Normalization*, las flechas curvas corresponden a *skip connections*, la letra T denota a una transposición, y D a capas densas. Las letras A, B, C y E corresponden a los hiperparámetros a optimizar, y representan el número de neuronas de la capa en cuestión. 27

Índice de cuadros

3.1.	Tabla obtenida de Valderrama <i>et al.</i> [75], presenta propuestas de uso de ML para el problema de estimación de potencial teórico. En inglés del original. Descripción original: <i>Several ML application proposals for the theoretical potential estimation. The input and target variables classification is presented in parenthesis.</i>	6
3.2.	Tabla obtenida de Valderrama <i>et al.</i> [75], presenta propuestas de uso de ML para el problema de estimación de potencial urbano. En inglés del original. Descripción original: <i>Several ML application proposals for the urban potential estimation. The input and target variables classification is presented in parenthesis.</i>	14
3.3.	Tabla obtenida de Valderrama <i>et al.</i> [75], presenta propuestas de uso de ML para el problema de estimación de potencial técnico. En inglés del original. Descripción original: <i>Several ML application proposals for the technical potential estimation. The input and target variables classification is presented in parenthesis.</i>	15
8.1.	Resultados de exploración de representaciones sobre validación utilizando 5-fold.	27
8.2.	Resultados de hiperparámetros óptimos.	28
8.3.	Resultados de redes optimizadas sobre test, en términos de la pérdida y el MAPE. Adicionalmente se presenta la dimensionalidad NDim de la representación interna de los datos en la última capa previo a la capa de <i>flatten</i> de cada arquitectura. Los intervalos de confianza consideran un 5% de incertidumbre.	29
8.4.	Presenta los tiempos promedio de entrenamiento y predicción de las distintas arquitecturas, en segundos, con intervalos de confianza sobre las 10 repeticiones con una significancia del 5%.	30
8.5.	Resultados de uso de representaciones intermedias de la red para ajuste de regresión lineal, con intervalos de confianza sobre las 10 repeticiones con una significancia del 5%.	30

Resumen

Los inminentes problemas asociados al cambio climático han acelerado los esfuerzos para lograr una pronta transición energética hacia fuentes renovables. Por ende, la estimación precisa y a gran escala del potencial de generación de distintas energías renovables toma una nueva importancia en el proceso de toma de decisiones y políticas públicas. La estimación fotovoltaica sobre las zonas urbanas es particularmente complicada dada la influencia de las oclusiones del sol por edificaciones adyacentes, la disponibilidad incierta de áreas favorables para las instalaciones y la ausencia de datos apropiados públicamente disponibles y confiables.

El Ministerio de Energía ha generado una base de datos del potencial fotovoltaico urbano de Santiago, que puede utilizarse para efectos de aprendizaje. Más aún, una segunda base de datos, obtenida desde el Servicio de Impuestos Internos, presenta descriptores de las edificaciones presentes en el territorio nacional, entre otras, el número de pisos o la superficie total construida. Esto permite consolidar una base de datos a nivel de manzana, la cual presenta una lista de edificios con sus respectivos descriptores y el potencial fotovoltaico, en términos del área favorable, de esta. Con estos datos se puede realizar aprendizaje automático para modelar la relación entre las características de las manzanas y su potencial. Esta prometedora aproximación tiene, sin embargo, una dificultad adicional: distintos registros (*i.e.* manzanas) tienen distinto número de edificaciones, y por ende distinta dimensionalidad. Esto requiere por lo tanto un manejo particular de los datos, los cuales tienen largo variable.

En el presente proyecto se propone aprovechar la capacidad de las redes convolucionales uno dimensionales de aprender patrones sobre secuencias de datos para realizar el aprendizaje sobre representaciones secuenciales de los datos disponibles. Más aún, se propone utilizar los datos tanto en la representación habitual “homogénea” (distintas posiciones de la secuencia corresponden a distintas instancias del mismo tipo de dato), como “heterogéneas” (distintas posiciones corresponden a distintos tipos de datos, en nuestro caso, distintos atributos de las edificaciones), como una tercera representación “mixta”, inicialmente homogénea seguida de heterogénea. Además, esto permitirá igualmente a la red aprender representaciones convolucionales significativas de los datos.

Si bien estas innovadoras representaciones heterogénea y mixta no permiten mejorar significativamente los resultados del estado del arte homogéneo, si se aprecian diferencias en los costos computacionales asociados, siendo la representación heterogénea significativamente menos costosa en tiempos de entrenamiento y predicción que la representación estándar homogénea. Esto se evidencia igualmente en el aprendizaje realizado sobre las representaciones intermedias, donde la representación heterogénea aprendida logra desempeños similares utilizando órdenes de magnitud menos dimensiones que las otras representaciones.

Palabras claves: Estimación Fotovoltaica Urbana, Redes Convolucionales, Representaciones Heterogéneas, Representaciones Homogéneas, Aprendizaje de Representaciones.

Abstract

The urgency of addressing climate change has accelerated the efforts to transition to renewable energy sources. Consequently, the precise and large-scale estimation of generation potential has gained importance for policy-making and decision-making processes. Among the challenges, estimating urban photovoltaic potential proves particularly difficult due to complex factors such as shadow projection from neighboring buildings, uncertain availability of favorable areas, and lack of reliable public data.

To aid in this estimation, the Ministry of Energy has created an urban photovoltaic potential database for Santiago, which can be utilized for learning purposes. Additionally, a second database obtained from the Internal Revenues Office provides building descriptors for the entire national territory, including information on the number of floors and total built area. By combining and georeferencing these datasets, a consolidated database is generated at the block level, containing building descriptors and photovoltaic potential in terms of favorable area. This data enables machine learning to model the relationship between block characteristics and their potential. However, a challenge arises from the varying number of buildings within different blocks, leading to different dimensionalities in the data, which requires specific handling.

This project proposes leveraging the capacity of one-dimensional convolutional neural networks to learn patterns from sequential representations of the available data. The data will be used in three representations: the standard “homogeneous” representation where different positions correspond to different instances of the same data, the “heterogeneous” representation where different positions represent different types of data (e.g., building attributes), and a “mixed” representation initially following the homogeneous pattern and transitioning to heterogeneous. This approach also allows the networks to learn meaningful convolutional representations of the data.

While the innovative heterogeneous and mixed representations do not significantly improve the results compared to the standard homogeneous representation, there are notable differences in computational costs. The heterogeneous representation proves to be considerably less computationally expensive in terms of training and prediction times. This is also evident in the learning achieved with intermediate representations, where the heterogeneous representation achieves comparable performance using significantly fewer dimensions compared to the other representations.

Key words: Urban Photovoltaic Estimation, Convolutional Neural Networks, Heterogeneous Representations, Homogeneous Representations, Representation Learning.

La inminencia de las peores consecuencias del cambio climático ha llevado a varios países a articular planes de transición energética acelerada, buscando la sostenibilidad y descarbonización de la matriz energética. Chile no ha quedado fuera de estas iniciativas, comprometiéndose, entre otros, mediante la firma y ratificación del acuerdo de París [68]. Este esfuerzo de gran envergadura requiere, por su naturaleza, una planificación de largo plazo, implementando políticas públicas e incentivos que permitan la transición energética de manera oportuna y sostenible. Por ende, la estimación precisa y a gran escala de los potenciales de generación de distintas energías renovables resulta un insumo importante para la toma de decisiones estratégicas.

Esta estimación conlleva diversos desafíos: puede considerar distintas tecnologías (como eólica, fotovoltaica, solar térmica, etc.), puede abarcar distintas áreas de aplicabilidad (es decir, urbana, rural) o incluso estimar distintos tipos de potenciales (potencial teórico, técnico, de mercado, etc.). En el caso fotovoltaico, Chile cuenta con bases de datos de radiación solar sobre todo el territorio nacional [57], sin embargo, en el paisaje urbano no existe ninguna caracterización más allá de estimaciones a nivel de ciudad. La rugosidad del terreno urbano conlleva dificultades suplementarias, tales como la necesidad de conocer la disponibilidad de áreas favorables para instalaciones o la proyección de sombras entre edificaciones adyacentes sobre estas áreas.

El Ministerio de Energía dispone de una base de datos de áreas favorables para efectos fotovoltaicos de las manzanas urbanas de Santiago, con alta granularidad espacial. Esta puede permitir aprender la relación entre descriptores de las edificaciones (obtenibles sobre todo el territorio nacional) y el potencial de generación fotovoltaica urbana.

El presente documento presenta una propuesta de exploración y comparación sistemática del desempeño de redes convolucionales uno dimensionales en la tarea de predecir el potencial fotovoltaico urbano sobre el conjunto de datos disponibles, aprendiendo representaciones convolucionales intermedias de largo fijo. Además, se compararán arquitecturas convolucionales 1 dimensionales con representaciones internas de los datos no solo homogéneas (distintas po-

siciones en las secuencias corresponden a diferentes instancias de la misma característica), si no también heterogéneas (distintas posiciones en las secuencias corresponden a diferentes características), y mixtas (distintas representaciones internas, tanto heterogéneas como homogéneas). En la medida de nuestro conocimiento, representaciones de naturaleza heterogénea o mixtas no han sido exploradas en la literatura.

En el resto de la propuesta, se presenta, primeramente, en el Capítulo 2, un marco general de la problemática de la transición energética y la relevancia de la estimación de potenciales de energías renovables, seguido de una descripción de distintas propuestas para abordar el problema de estimación fotovoltaica, tanto urbanas como no urbanas, usando *Machine Learning*, en el Capítulo 3. Luego, el Capítulo 4 entrega una visión general del caso chileno, discutiendo la realidad de la estimación fotovoltaica nacional y los datos asociados disponibles. Seguidamente, presentaremos una formulación teórica del problema de aprendizaje asociado en el Capítulo 5, mientras el Capítulo 6 presenta las características particulares de las redes convolucionales uno dimensionales. La propuesta de tesis se presenta en el Capítulo 7, discutiendo la motivación que guía esta, y la relevancia de las distintas representaciones de los datos, presentando las hipótesis y los objetivos específicos. Seguidamente, el Capítulo 8 detalla el conjunto de experimentos que se han realizado, y sus resultados. Finalmente, el Capítulo 10 presenta las conclusiones derivadas del presente trabajo, mientras el Capítulo 11 presenta eventuales trabajos futuros que surgen de lo observado durante este proceso.

Existe un amplio consenso multidisciplinario llamando a la descarbonización de la matriz energética para paliar las peores consecuencias del cambio climático [2]. Cambios tecnológicos están posibilitando esta transición, en la cual el paisaje urbano jugará un rol importante en las próximas décadas [74], tanto por el lado del consumo, dado el crecimiento urbano global, como por el lado de la generación, donde la generación fotovoltaica distribuida puede contribuir significativamente a la descarbonización. Es por esto por lo que la estimación confiable y a gran escala del potencial de generación fotovoltaico de zonas urbanas es un insumo esencial para facilitar y optimizar la toma de decisiones enfocadas a guiar esta transición [9]. Esta estimación presenta diversos desafíos, y siguiendo la conceptualización utilizada por varios autores [9, 36, 81], forma parte de un proceso de estimación jerárquico compuesto de a lo menos 5 pasos sucesivos. A saber, los 5 siguientes potenciales son descritos:

1. **Potencial Teórico o Físico:** Corresponderá a la energía total entregada por el recurso solar en un área determinada. Esta estimación suele involucrar procesos físicos ampliamente estudiados, teledetección satelital y validación con mediciones *in situ*.
2. **Potencial Geográfico o Urbano:** Consistirá del potencial teórico de energía recibido sobre las áreas donde efectivamente puede ser capturada esta energía (por ejemplo, los techos urbanos), considerando la inclinación del área disponible. Además, deberá tomar en cuenta las proyecciones de sombras sobre dichas áreas, en particular de las características topográficas en el caso geográfico, mientras en el caso urbano corresponderá a sombras de edificaciones adyacentes.
3. **Potencial Técnico:** Se dará por el potencial energético que puede ser generado por un sistema fotovoltaico real, instalado en el área disponible determinada. Este se calcula a partir de ecuaciones físicas, incorporando las eficiencias y pérdidas durante el proceso de transformación de energía solar a energía eléctrica mediante el panel y el inversor.
4. **Potencial Económico:** Vendrá dado por la factibilidad económica de efectivamente instalar paneles en la determinada área. Debe considerar el costo tecnológico, de instalación

y mantención, así como el costo de la energía en la zona de participación.

5. Potencial de Mercado: Este último corresponderá al potencial realizable de un área dada las restricciones y fricciones de mercado, ya sea por componente socioculturales (*e.g.* reticencia a nuevas tecnologías), regulaciones o políticas públicas relevantes.

Así, el problema de la estimación de potencial fotovoltaico es un problema transdisciplinario. La jerarquía presentada acota los distintos problemas y los transforma en sucesivas estimaciones, las cuales en algunos casos pueden realizarse independientemente. Dado que distintas estimaciones corresponden a fenómenos de distinto carácter, conviene aplicar métodos apropiados para la naturaleza de cada caso. En efecto, Assouline *et al.* [9] han clasificado los distintos modelos utilizados a lo largo de esta estimación, en el estado del arte, en las siguientes familias de modelos: físicos o empíricos, geoestadísticos, valor constante, muestreo, sistemas de información geográfica (GIS) o LIDAR, y finalmente modelos de machine learning (ML).

Cada una de estas metodologías tiene distintas ventajas y desventajas, y en algunos casos la disponibilidad de datos, tiempo y recursos determinará cual es la mejor alternativa para la estimación que se desee realizar. En nuestro caso, este corresponderá a modelos de ML, dado que en el caso urbano los modelos físicos son en la práctica intratables; modelos geoestadísticos requieren data dispersa espacialmente de la cual no contamos, y, por lo demás, suelen ajustarse mejor a realidades físicas; métodos de valor constante y muestreo tienen malas precisiones; mientras métodos GIS o LIDAR requieren de inversiones mayores, y en el caso particular de LIDAR pueden requerir tiempos prohibitivos o incluso problemas regulatorios (*e.g.*, permisos de vuelos de drones sobre áreas urbanas).

Estado del Arte

En este Capítulo discutiremos algunas propuestas de la literatura que abordan distintos pasos de la estimación de potencial fotovoltaica. No se presentarán ningún estudio del potencial de mercado, ni tampoco nos enfocaremos en las particularidades de las estimaciones de potencial económico dado que estos se alejan mucho del ámbito de la presente propuesta. Para mayor detalle y variedad de propuestas se puede revisar la publicación por Valderrama *et al.* [75], realizada en el contexto de esta iniciativa. Al final de este capítulo se presenta las tablas de resumen de los manuscritos para la estimación de potencial teórico, técnico y urbano revisados y clasificados en el contexto de la publicación referenciada, los cuales van más allá de las publicaciones descritas en el presente capítulo.

Mainzer *et al.* [49] describen una metodología para la estimación del potencial fotovoltaico, incluyendo los potenciales teóricos, urbanos, técnicos y económicos. Los autores utilizan diversas fuentes de datos, como huellas de los edificios (*building footprints*), imágenes satelitales de libre disposición, telemetría de radiatividad, entre otras. En su estimación del potencial teórico utilizan un modelo determinista estándar de irradiancia solar (modelo físico). El potencial urbano lo estiman combinando diversas técnicas, incluyendo algoritmos de procesamiento de imágenes (estándar en software GIS) para detección de bordes o techos, estimación de la inclinación de los techos mediante aleatorización de estos (gaussiana con media y dispersión acoplada a mediciones previas, por ende, método de muestreo) o la detección de instalaciones fotovoltaicas preexistentes sobre área disponible (versión de “Alexnet” [45] sobre imágenes satelitales). Los autores posteriormente comparan los resultados de su estimación con un modelo puramente físico, obtenido mediante una representación urbana tridimensional preexistente e imágenes clasificadas manualmente para el caso de paneles preexistentes. Mainzer *et al.* evalúan el desempeño de sus modelos mediante la precisión del cálculo del azimut (desviación angular respecto al norte) de la línea central de los techos obtenidos mediante metodología GIS, reportando un 70% de precisión; y calculando la precisión de la clasificación de techos con y sin preexistencias, alcanzando un 90% en este caso.

En [34], Anwar y Khatib proponen un modelo híbrido para la predicción de la radiación

Spatial Scale	Temp. Res.	ML Methods	Target Value(s)	Inputs
Whole Country	Monthly	RF	Weather values, horizontal irradiance (<i>Environmental</i>)	DEM, Building topologies (<i>Environmental</i>)
Whole Country	Monthly	ANN	Solar Irradiance (<i>Environmental</i>)	Lat, Lon, month; average of: temp, RH, sunshine duration, windspeed, precipitation (<i>Environmental</i>)
26 Cities	Monthly	ANN	Mean Solar Radiation (<i>Environmental</i>)	Temperature, sunshine duration, RH, precipitation (<i>Environmental</i>)
Single City	Monthly	ANN	Global Solar Radiation (<i>Environmental</i>)	Average, max, and min of temperature; Lat, Lon, extraterrestrial global solar radiation, clearness index, altitude, sunshine hours (<i>Environmental</i>)
Single City	Seasonal	GB	Solar irradiation over urban surfaces and terrains (<i>Environmental</i>)	Parametric irradiance data, sky view factor and surface orientation (<i>Environmental</i>)
195 Cities	Monthly	ANN	Solar Irradiation Prediction (<i>Environmental</i>)	Lat, Lon, altitude, month; average: sunshine duration, temperature, RH (<i>Environmental</i>)
15 Cities	Monthly	ANN, Genetic Algorithm	Average solar radiation (<i>Environmental</i>)	Lat, Lon, altitude; average: fraction of sunshine, cloud amount, sunshine duration, temperature, RH, wind speed (<i>Environmental</i>)
Whole country	Daily	ANN, Ridge Regression	Global solar irradiance (<i>Environmental</i>)	Temperature, precipitation (<i>Environmental</i>)
Whole Country	Hourly	ELM Ensemble	Global solar irradiance (<i>Environmental</i>)	Lat, Lon, altitude, hour, coarse-resolution satellite measurements (<i>Environmental</i>)
Several Points	Daily	ANN	Solar Irradiation (<i>Environmental</i>)	Day of year, clearness index, extraterrestrial radiation (<i>Environmental</i>)
13 Points	Daily, Hourly	FFANN	Solar Global Radiation Estimation (<i>Environmental</i>)	Lat, Lon, altitude, month, time, air temperature, wind speed, RH, rainfall (<i>Environmental</i>)
41 Points	Monthly	ANN	Ratio of mean daily global solar radiation divided by extraterrestrial radiation (<i>Environmental</i>)	Lat, Lon, altitude, sunshine duration (<i>Environmental</i>)
12 Points	Monthly	ANN	Solar Radiation (<i>Environmental</i>)	Lat, Lon, altitude, month, mean sunshine duration, mean temperature (<i>Environmental</i>)
60 Met. Stations	Monthly	ANN	Global Solar Radiation (<i>Environmental</i>)	Lat, Lon, altitude (<i>Environmental</i>)
17 Met. Stations	Monthly	ANN	Solar Radiation (<i>Environmental</i>)	Lat, Lon, altitude, month, mean sunshine duration, mean temperature (<i>Environmental</i>)
Single Site	Hourly	RF+FA, FFANN, FFANN+FA	Global Solar Radiation (<i>Environmental</i>)	Sunshine ratio, RH, temperature, month number, day number, hour number (<i>Environmental</i>)
3 Sites	Daily	RF	Solar Radiation (<i>Environmental</i>)	Sunshine hours, maximum possible sunshine hours (<i>Environmental</i>)
Single City	Hourly	Deep ANN	Hourly solar radiation (<i>Environmental</i>)	Stationary and high resolution Satellite Imagery (<i>Remote Sensing</i>)

Tabla 3.1: Tabla obtenida de Valderrama *et al.* [75], presenta propuestas de uso de ML para el problema de estimación de potencial teórico. En inglés del original. Descripción original: *Several ML application proposals for the theoretical potential estimation. The input and target variables classification is presented in parenthesis.*

solar por hora. En particular, los autores tienen como objetivo predecir los valores horarios para los datos medidos en un único sitio en Malasia. Las variables de entrada utilizadas por los autores son: la proporción de horas de sol, la humedad relativa, la temperatura, el número del mes actual, el número del día actual y el número de la hora. El método de aprendizaje automático utilizado por los autores es el modelo Random Forest (RF) [16]. Los modelos RF buscan mitigar la tendencia de los modelos de árbol de regresión tradicionales a sobreajustar los datos disponibles [26]. Esto se logra mediante un conjunto de varios árboles y el uso no de una, sino de dos formas de bagging [15]. Cada uno de los árboles del modelo se ajusta a una muestra bootstrap de los datos, utilizando solo una muestra bootstrap de las variables disponibles. No obstante, estos modelos poseen varios hiperparámetros, como el número de árboles o el número máximo de hojas por árbol, que deben ajustarse para obtener un rendimiento óptimo. Los autores proponen utilizar el algoritmo metaheurístico *firefly algorithm* (FA) [84, 85] para el ajuste de los hiperparámetros. Inspirado en el comportamiento de apareamiento de las luciérnagas en entornos naturales, el FA es una técnica metaheurística de optimización basada en población y en la naturaleza. En cada iteración, cada solución se representa por un punto en el espacio. Su brillo es proporcional al rendimiento de la solución. Al final de cada iteración, las soluciones pueden acercarse entre sí siguiendo leyes de atracción determinadas por su posición relativa y brillo. Utilizando este algoritmo, los autores optimizan su modelo de RF, comparando sus resultados con la red neuronal artificial de *Feed Forward* (FFANN) presentada en [41] y otra FFANN con el número de neuronas optimizado utilizando FA igualmente. El mejor rendimiento encontrado por los autores corresponde al RF optimizado con FA, logrando un error porcentual absoluto medio (MAPE) de 6.3826 %.

Otra propuesta, por Benmouiza & Cheknane [12], utiliza modelos tipo ARMA y redes neuronales recurrentes para predecir la radiación solar global con un día de antelación. Los autores utilizan estos modelos estadísticos y de ML para estudiar el problema con una alta resolución temporal (de 30 o 1 segundo), para un punto de medición particular. Por ende, su estudio corresponde a una estimación puntal del potencial teórico, con altísima resolución temporal. Este tipo de estimaciones resulta valioso para planificación de corto plazo (*e.g.* coordinador eléctrico y operador de una planta). Los autores reportan un error cuadrático medio normalizado (NRMSE) de 0,2034 en el mejor de sus modelos.

Por otro lado, Voyant *et al.* [21] presentan modelos de series de tiempo de radiación solar global para predicciones de esta a gran escala. En particular, ellos buscan predecir la radiación solar global con una hora de antelación sobre la isla de Córcega, es decir, estimación del potencial teórico. Los autores utilizan datos derivados de imágenes satelitales con resolución espacial de 2,5 km, modelando así el recurso solar como serie de tiempo temporal 2-dimensional. Sobre cada uno de los puntos espaciales de esta serie de tiempo, los autores entrenan una red neuronal densa. Evaluando el desempeño predictivo de cada una de estas redes neuronales sobre su serie de tiempo correspondiente, los autores reportan un promedio sobre el territorio en métrica NRMSE de 16,6.

En la propuesta realizada por Reddy y Ranjan [62], los autores utilizaron datos medidos de radiación global de once estaciones diferentes para entrenar una red neuronal artificial para realizar predicciones horarias, validando los resultados en otras dos estaciones. Como entrada para su red, los autores utilizan la latitud, longitud, altitud, número de mes, hora del día, temperatura del aire, velocidad del viento, humedad relativa y presencia de lluvia.

La red propuesta es una FFANN, densamente conectada con dos capas ocultas, la primera compuesta por 8 neuronas y la segunda compuesta por 7 neuronas. Los autores entrenan diferentes redes para diferentes regiones (norte o sur de la India), y diferentes estaciones climáticas (verano, lluviosa e invierno) utilizando el algoritmo de *backpropagation* [64] para el entrenamiento. Utilizando estas redes, los autores informan una desviación relativa media absoluta del 4.07 %, lo que proporciona una estimación precisa y detallada de la radiación para áreas sin equipos de medición. Sin embargo, la elección de especificar diferentes redes para diferentes regiones espaciales restringe la generalización espacial del método, ya que estas redes necesitarían ser entrenadas con nuevos datos para predecir valores de radiación en regiones distintas a las ya mencionadas.

Otro trabajo con características similares es presentado por Rumbayan *et al.* en [66]. En este trabajo, los autores utilizan una FFANN con una capa oculta para predecir la radiación solar promedio mensual sobre Indonesia. Como entradas, los autores proporcionan a la red varias variables, a saber: promedios de temperatura, humedad relativa, duración de la luz solar, velocidad del viento y precipitación; la latitud, longitud, altitud y el número del mes. Los autores optimizan la arquitectura de la red modificando el número de neuronas en la capa oculta entre 4 y 15, encontrando el mejor error para 11 neuronas, pero señalando que todas las arquitecturas presentan un MAPE por debajo del 5 %. En cambio, la mejor configuración de la red logra un MAPE de solo el 3.29 %. Los autores utilizan esta red para mapear la radiación solar mensual disponible en el resto del país, informando un alto potencial de radiación en Indonesia.

Al estimar el potencial urbano, distintas consideraciones surgen en función de las prioridades de los autores. Por ejemplo, para estimar el área potencial disponible en los techos para instalaciones de paneles solares fotovoltaicos (PV), se deben tener en cuenta las formas de los tejados. Las inclinaciones de las cubiertas afectan la cantidad de radiación recibida, y los paneles PV no se ajustarán a todos los tejados según sus dimensiones particulares. En este contexto, Mohajeri *et al.* presentan en [54] una metodología de aprendizaje automático (ML) para la clasificación de las formas de los tejados en Ginebra, Suiza. El modelo de ML utilizado corresponde a la Máquina de Vectores de Soporte (SVM) [20] utilizando funciones de base radial. Las SVM aprenden una proyección implícita de alta dimensionalidad de los datos de entrada, en un espacio que separa linealmente las diferentes clases, lo cual se utiliza para un esquema de clasificación mediante un hiperplano separador. Los autores clasificaron manualmente 717 edificios (alrededor del 6 % del total del conjunto de datos) en 6 clases finales de formas de tejado, utilizando Google Earth y un mapa aéreo de alta calidad [1], que sirven como datos de entrenamiento. Se utilizan un total de 35 variables como entradas para el algoritmo, que incluyen el número de superficies de cada edificio, las distribuciones de pendientes, el porcentaje de espacio en el tejado para cada pendiente, entre otros. Los autores informan una precisión promedio del 66 % en su conjunto de validación. Utilizando los resultados de este método, los autores refinan aún más la clasificación de las cubiertas según su área potencial para instalaciones de PV y el potencial de energía solar. Con esa información, los autores logran presentar el área potencial promediada y el potencial de energía PV para cada tipo de forma.

Otra consideración importante en algunos espacios urbanos es la presencia de arreglos fotovoltaicos preexistentes en áreas disponibles. Cuantificar esto es útil tanto para evaluar el

potencial restante de PV como para medir la adopción de PV. En [50], Malof *et al.* combinan una Red Neuronal Convolutiva (CNN) profunda [46] y un modelo de RF para la detección de arreglos PV previos en entornos urbanos. Para este propósito, los autores utilizan imágenes aéreas ortorrectificadas [89] con bandas RGB de alta resolución disponibles públicamente [14], clasificando manualmente 2700 ubicaciones de arreglos PV. Los autores entrenan el modelo de RF para asignar una probabilidad de que el píxel corresponda a un array de sistema PV. Por otro lado, la red CNN, compuesta por 6 capas convolucionales con 3 capas de *max pool* y una capa densa final, se entrena en parches de píxeles de 40x40. A esta red se le asigna la tarea de clasificar el parche como conteniendo sistemas PV o no. Para la evaluación, los autores proponen un esquema de clasificación de dos etapas: primero, se identifican los parches con presencia probable de PV mediante el RF, para luego ser clasificados por la CNN. Los autores informan las curvas de *precision-recall* [17], y una precisión de identificación de paneles del 80 % para el esquema de clasificación de dos etapas.

También es necesario abordar el problema de identificar y cuantificar la superficie de los techos en un área. Una propuesta para resolver este problema es presentada por Huang *et al.* [31]. Los autores implementan un modelo U-Net [65], un tipo particular de CNN para segmentación de imágenes que, a través de su arquitectura específica, puede incorporar características altamente sofisticadas de la imagen con características locales más finas y relevantes. Para entrenar la red, los autores utilizan 180 imágenes de alta resolución de una base de datos disponible públicamente [48] de varias ciudades diferentes. Utilizando la red entrenada, los autores la aplican a la ciudad de Wuhan, China. Para validar sus resultados, los autores estiman primero un valor de generación de energía utilizando los resultados de la red. Los autores logran esto transformando el área disponible estimada en una producción de energía mediante el establecimiento de un valor de radiación horizontal global, y utilizando una tasa de conversión PV estándar constante. Como referencia de comparación, los autores utilizan la modelización 3D de los edificios y la simulación de la radiación incidente sobre un área urbana más reducida. Sobre esa área, los autores informan una desviación del 9.51 % entre ambos modelos. Los autores destacan la velocidad de predicción del modelo de ML en comparación con el esquema de modelización 3D.

Otros atributos pueden resultar relevantes para el problema de estimación urbana igualmente, como el uso del suelo. Este es el caso de [86], donde Zhang *et al.* proponen un método para estimar no solo el potencial de energía solar fotovoltaica en los tejados, sino también el tipo de uso del suelo. Al igual que Huang *et al.*, los autores utilizan una red U-Net para estimar el potencial de energía solar fotovoltaica en los tejados de Wuhan, China. Sin embargo, también implementan otra arquitectura convolutiva, la red VGG16 [67], para la predicción del uso del suelo. Con esta red, los autores implementan un esquema de transferencia de aprendizaje [73]: preentrenan la red en un gran conjunto de datos de uso del suelo de ciudades europeas e imágenes satelitales. Los autores evalúan el uso de la red preentrenada para predecir el valor objetivo del uso del suelo en Wuhan, sin embargo, encuentran un bajo rendimiento que se atribuye a diferentes patrones del uso de suelo. Por lo tanto, los autores continúan entrenando la red con un menor número de ejemplos clasificados de la ciudad de Wuhan, reportando una precisión final del 81.45 % en el conjunto de pruebas de Wuhan.

Otros trabajos se centran en estimar directamente la radiación solar en superficies urbanas, considerando, cuando es posible, las variaciones temporales. Por ejemplo, en [77],

Vartholomaios presenta un algoritmo de *gradient boosting* (GB) para la predicción de la radiación solar estacional en superficies urbanas. El autor modela paramétricamente los datos de irradiación utilizando el simulador de iluminación por trazado de rayos RADIANCE. Los factores que afectan la estimación son el factor de vista al cielo (sky view factor, que representa la proporción del área despejada del cielo) [19] y las orientaciones de las superficies en términos de ángulo horizontal y cenital. El autor optimiza varios hiperparámetros del algoritmo GB, como el número de estimadores, la profundidad máxima de cada árbol, el número mínimo de muestras para una división y para una hoja, y la fracción de submuestreo para cada iteración del modelo. El autor informa que la mejor combinación de hiperparámetros logra un MAPE de solo el 3.4 % en las pruebas. Sin embargo, al comparar los datos predichos con datos del mundo real, Vartholomaios encuentra resultados más pobres, con un MAPE promedio del 16 % en Heliópolis, Atenas, Grecia. Esta discrepancia se atribuye a una suposición subyacente errónea de que el factor de vista al cielo por sí solo puede modelar con precisión la cantidad de radiación incidente.

Otro trabajo realizado por Assouline *et al.* [7], presenta un método basado en máquinas de soporte vectorial SVR (*support vector regression*), para estimar la superficie de tejados disponibles, su ángulo inclinado promedio y el coeficiente de sombreado. Las variables de entrada utilizadas son la densidad de edificios, la densidad de población, el área promedio de huella del edificio, el área total de huella del edificio y algunas tipologías de edificios, como el período de construcción, el número de pisos o el recurso energético principal utilizado. Estas variables de entrada se agregan a nivel de municipio. Los datos de entrenamiento corresponden a 46 municipios etiquetados. En términos de la raíz del error cuadrático medio (RMSE), los autores informan un valor de 0.0154 para la relación entre la superficie de tejado disponible y el área de la huella del edificio, 0.54 para el ángulo de inclinación (en grados) y 2.14 para el coeficiente de sombreado (en una escala de 0 a 255). Si bien la superficie de tejado disponible y el ángulo de inclinación son variables atemporales, los autores estiman un valor mensual para el coeficiente de sombreado. Esto es valioso ya que la posición del sol varía a lo largo del año, lo que puede afectar la cantidad de luz solar recibida y, por lo tanto, el potencial de generación de energía. También vale la pena mencionar que los autores utilizaron enfoques de ML similares para una estimación del potencial teórico en términos de la radiación horizontal global, la radiación horizontal difusa y la radiación directa normal.

También en Suiza, Walch *et al.* presentan en [80] un método de ML para estimar el espacio de tejado disponible para instalaciones fotovoltaicas sin eliminar las superestructuras previamente existentes, y mejorar la fracción de sombreado del tejado a nivel horario. Esta variable se modela sobre un modelo 3D de baja resolución ($2 \times 2 \text{ m}^2$), mientras que los datos de entrenamiento objetivo se obtienen de un modelo 3D más preciso ($0.5 \times 0.5 \text{ m}^2$). Para estimar estas variables, los autores entrenan modelos RF sobre el cantón de Ginebra, donde se dispone de datos. Como características de entrada, los autores incluyen el área inclinada, el ángulo de inclinación, el perímetro, el tipo de edificio, el número de pisos o el período de construcción, entre otras variables. Para este modelo, los autores informan un RMSE de 0.12. Los autores también informan un RMSE de 0.18 para el modelo RF del factor de sombreado. En este caso, los autores también realizaron una estimación teórica utilizando conjuntos de ELM (*extreme learning machines*).

Igualmente, algunas propuestas evalúan el potencial técnico. Si bien en la actualidad

existen fórmulas precisas y rápidas para la relación entre la radiación incidente y el estado del arreglo de paneles solares y su producción de energía, estos no siempre están disponibles, debido a las diferencias entre sistemas y materiales. Algunas propuestas utilizaron el aprendizaje automático para modelar la relación subyacente sin necesidad de un conocimiento previo extenso del sistema de paneles solares en particular, como se muestra en [60] de Premrudeepreechacham y Patanapirom. Los autores implementan y comparan dos tipos de redes neuronales, en concreto la arquitectura *feed-forward* y la red neuronal de función de base radial (RBF ANN) [22]. Utilizando como entradas la radiación y la temperatura, los autores pretenden modelar, por un lado, el problema del seguimiento del punto de máxima potencia (prediciendo la corriente y el voltaje máximos del sistema dadas las condiciones). Por otro lado, con la inclusión del voltaje de carga del sistema, pretenden modelar la corriente de carga. Para el problema del seguimiento del punto de máxima potencia, las arquitecturas utilizadas corresponden a una red con 150 neuronas ocultas, mientras que para el modelado de la corriente se utilizan redes con 100 neuronas ocultas (tanto para FFANN como para RBF ANN). Los autores informan buenos resultados para ambos problemas y ambas arquitecturas de red, encontrando una desviación de menos del 1 % para casi todas las configuraciones del sistema de paneles solares, excepto para la corriente máxima predicha por FFANN, donde el 99 % de las configuraciones presentaron una desviación inferior al 1 %.

Otra propuesta que utiliza FFANN para el problema de estimación técnica corresponde a Mellit *et al.* [52]. Aquí, los autores proponen predecir los parámetros óptimos de dimensionamiento de un sistema PV y de almacenamiento utilizando solo la latitud y la longitud del sistema. Esto podría permitir, en teoría, una estimación a gran escala y configurada de manera óptima del potencial PV al considerar sistemas PV y de almacenamiento. Los autores entrenan una FFANN de 1 capa oculta con 6 neuronas utilizando 50 puntos de datos diferentes, dejando 4 puntos para las pruebas. Todos los puntos de datos se encuentran en Argelia. Los autores informan errores entre el 2.38 y el 6.1 por ciento, concluyendo que la red presenta un comportamiento aceptable. Los autores también destacan la rapidez de este método de estimación, especialmente una vez que la red está entrenada, en comparación con los enfoques de simulación de la época.

Otras propuestas buscan abordar diferentes configuraciones de sistemas PV, que son más difíciles de modelar o aún no se han explorado completamente. Por ejemplo, Ghani y Carson presentan en [27] una FFANN que puede modelar la eficiencia de conversión de un arreglo PV y colector térmico. Este tipo de sistema corresponde a aquel que incluye tanto un arreglo PV como un sistema colector térmico basado en líquido. Aquí, el rendimiento del PV se mejora mediante la refrigeración del arreglo a través del flujo de líquido que extrae calor del sistema. Los autores entrenan la red utilizando 288 simulaciones numéricas realizadas previamente utilizando un método validado. Las variables de entrada corresponden a la relación de flujo de masa (es decir, la cantidad de fluido que se mueve a través del sistema durante un período dado) y la relación de aspecto del arreglo. Se entrenan simultáneamente cuatro objetivos diferentes, que corresponden a diferentes configuraciones del flujo permitido y la dirección del fluido. Los autores informan un valor R de 0.97916 para el conjunto de pruebas. Los autores señalan que una de las principales ventajas de los métodos de ML en comparación con la modelización por simulación es que, una vez que los modelos están entrenados, los tiempos de predicción son muy rápidos. Por lo tanto, se pueden realizar optimizaciones exhaustivas

de las características de los sistemas PV en períodos cortos.

Otros autores han explorado el uso de ML para la predicción de series temporales PV a corto plazo. Si bien esto no estima el potencial PV total, ayuda a optimizar la infraestructura PV existente y futura, lo que permite una mejor gestión de la energía y, potencialmente, una mayor potencia total de instalación PV. Un ejemplo de estos métodos se propone por Zhou *et al.* en [90]. En su trabajo, los autores desarrollan un conjunto de redes LSTM (Memoria a Corto y Largo Plazo) [30] con un mecanismo de atención [78, 61]. El modelo propuesto utiliza la serie temporal de temperatura y la serie temporal de potencia previa para predecir el siguiente paso de la potencia de salida. Las resoluciones temporales utilizadas por los autores son de 7.5, 15, 30 y 60 minutos. Los autores entrenan dos redes LSTM con mecanismos de atención, una en la serie de potencia y otra en la serie de temperatura. Finalmente, las salidas de ambas redes se fusionan y se alimentan a una FFANN que se entrena para realizar la predicción final. Los autores informan un MAPE final de 24.65 % para la serie temporal con resolución de 7.5 minutos. Los autores comparan estos resultados con varios otros modelos, como el modelo de persistencia, el Modelo de Promedio Móvil Integrado Auto-regresivo con Variable Exógena [59], una FFANN de una sola capa y una LSTM sin mecanismo de atención, informando que en todos los casos, el método de conjunto propuesto supera a los métodos de referencia. Aquí, tanto las variables de entrada como las variables objetivo corresponden al dominio ambiental.

Si nos enfocamos en la realidad chilena, los autores Campos *et al.* [18] presentan una metodología de estimación del potencial de generación distribuido en las zonas urbanas del país. Para esto, los autores realizan una detallada estimación para el caso particular de la ciudad de Concepción. Primeramente, utilizan una base de datos de radiación solar de 1 año con resolución horaria para el potencial teórico. Para el potencial urbano en cambio, los autores simulan una ciudad tridimensional a partir de los permisos de construcción del área urbana, sobre la cual calculan proyecciones de sombra y estiman área de techos disponibles. Por otro lado, estiman el potencial técnico con métodos de valor constante estándar, para finalmente estimar el potencial económico utilizando el indicador LCOE (Levelized Cost of Energy). Con esta metodología los autores encuentran un valor que estiman bastante preciso sobre la ciudad de Concepción, lo cual luego generalizan al resto del país. Sin embargo, dada la complejidad (tanto computacional, como de ejecución, y de levantamiento de datos) de su metodología, los autores suponen relaciones lineales entre el área de techos total y áreas disponibles con el número de habitantes de la ciudad, entregando un valor estimativo para cada ciudad de Chile. Los autores no son capaces de presentar alguna comparación de sus resultados dada la ausencia de datos previos en el territorio.

Otra propuesta que igualmente se hace cargo de la realidad urbana, por Assouline *et al.* [10], estima el potencial teórico, urbano y técnico sobre los techos de Suiza. Los autores realizan una estimación del valor promedio mensual de generación, durante un año, con una resolución espacial de 200 m utilizando datos de diversas fuentes, entre los cuales podemos destacar registros gubernamentales de edificaciones (ya sea mediante descriptores de edificaciones o modelos tridimensionales LIDAR para algunas áreas), data vectorial de clúster de edificaciones para identificación de zonas urbanas, modelos tridimensionales de techos de las edificaciones de igual forma que de superestructuras en techos. Los autores utilizan modelos de ML, en particular entrenando diversas instancias de random forest para predecir variables

climáticas (temperatura, nubosidad, precipitaciones y fracción de horas soleadas a partir de la latitud, longitud y altura), la radiación solar horizontal, el área disponible sobre los techos y el área de sombra sobre estos, igualmente como la insolación de las áreas con sombras parciales. Complementan estos modelos utilizando estimaciones de valor constante, en particular para la eficiencia de los paneles fotovoltaicos y las pérdidas de estos. Igualmente, en algunos casos los autores se apoyan de algoritmos de índole GIS para la extracción de datos geométricos de los techos disponibles. Los autores logran comparar sus resultados sobre un sampling de mediciones en un área particular de Suiza donde un proyecto anterior ha sido realizado, obteniendo un promedio de NRMSE entre ambos de 26.

Adicionalmente a las publicaciones presentadas en detalle previamente, se presentan tablas obtenidas de una publicación de Valderrama *et al.* [75] asociada a la presente iniciativa. En estas se resumen aproximaciones utilizando ML para resolver el problema de los potenciales teóricos, técnicos y urbano.

De la diversidad de trabajos presentados, rescatamos la heterogeneidad de las fuentes de datos utilizadas para las estimaciones, y como estas habilitan las distintas estimaciones y métodos a utilizar. Por ende, las metodologías deben ser *ad hoc* a cada territorio, dada las características de los datos disponibles sobre este.

Spatial Scale	Temp. Res.	ML Methods	Target Value(s)	Inputs
Whole Country	Monthly	RF	Weather values, irradiance, available rooftops, shaded surface proportion, shadow proportion (Environmental)	DEM, Building topologies (Morphological)
Several Muni.	Monthly	SVR	Weather variables, radiation values, rooftop radiation (Environmental)	Building descriptors, Lat, Lon, altitude (Morphological, Environmental)
Single City	Atemporal	SVM	Roof shape classification (Morphological)	Building descriptors such as aspect ratio, slope distributions (Morphological)
Single City	Hourly	FFANN	Horizontal radiation over roof surfaces (Environmental)	Day of year, hour of day (Environmental)
Several Muni.	Monthly	SVR	Average available rooftop surface, angle, shading coefficient (Morphological)	Population and building density, land use data, building typologies (Morphological, Sociocultural)
National	Monthly	SVM, RF	Monthly solar radiation, available rooftops, rooftop characteristics, shading factor (Environmental, Morphological)	DEM, land cover data, building descriptor data (Morphological)
Several Cities	Annual	SVR, RF, LinReg, KNN, EML Ensemble	Solar irradiation on buildings roofs (Environmental)	Rooftop characteristics, DSM derivatives, annual horizontal irradiation (Environmental, Morphological)
9.6 million rooftops	Hourly	RF, ELM Ensemble	Extrapolate effect of roof superstructures, enhance shading data (Environmental)	DEM, roof superstructure information (Morphological)
District	Atemporal	K-Means	Potential PV installation clusterings (Morphological)	Horizontal insolation, module slope, orientation (Environmental, Morphological)
Single City	Atemporal	CNN	Previous rooftop PV installation (Morphological)	Aerial Imagery (Remote Sensing)
Multiple Districts	Atemporal	SVM, FFANN	Rooftop detection (Morphological)	Satellite imagery (Remote Sensing)
Single City	Atemporal	Deep ANN	Rooftop detection (Morphological)	Satellite imagery, LIDAR, building typology (Remote Sensing)
Single City	Atemporal	SVM	Roof shape classification (Morphological)	LIDAR 3d data (Remote Sensing)
Single City	Monthly	LinReg	Urban PV potential (Morphological)	LIDAR 3d buildings, solar irradiation and PV models (Remote Sensing, Environmental)
National	Atemporal	SVM	Urban potential variation (Morphological)	Population, number of buildings, groundfloor area, street length (Morphological, Sociocultural)
Single City	Seasonal	GB	Solar irradiation over urban surfaces and terrains (Environmental)	Parametric irradiance data, sky view factor, surface orientation (Morphological, Environmental)
135 km ²	Atemporal	RF, CNN	PV array detection (Morphological)	Orthorectified RGB 0.3m resolution aerial photography (Remote Sensing)
Single City	Atemporal	CNN, Transfer Learning	Identification of land use type, rooftop solar potential (Morphological)	Satellite imagery, land use type (Remote Sensing, Sociocultural)
71 buildings	Atemporal	CNN	Roof segment and superstructure segmentation (Morphological)	Building footprints, aerial imagery (Remote Sensing, Morphological)
Single City	Atemporal	Deep ANN	Identification of rooftop surface availability (Morphological)	Satellite imagery (Remote Sensing)
Single City	Atemporal	Deep ANN	Rooftop shape categorization (Morphological)	Satellite imagery (Remote Sensing)
Single City	Atemporal	Deep ANN	Rooftop area estimation (Morphological)	Satellite imagery, land use (Remote Sensing, Sociocultural)
Single City	Atemporal	Deep ANN	Noise barrier identification (Morphological)	Street view open source imagery (Remote Sensing)

Tabla 3.2: Tabla obtenida de Valderrama *et al.* [75], presenta propuestas de uso de ML para el problema de estimación de potencial urbano. En inglés del original. Descripción original: *Several ML application proposals for the urban potential estimation. The input and target variables classification is presented in parenthesis.*

Spatial Scale	Temp. Res.	ML Methods	Target Value(s)	Inputs
National	Hourly	ANN, LinReg	PV electric output (Environmental)	Radiation, weather data, PV system temperature (Environmental) [40]
Aspatial	Atemporal	FFANN	Photovoltaic conversion efficiency of a building integrated PV/thermal collector array (Specifications)	Array shape, mass flow rate (Environmental) [27]
Single point	15 Minutes	SVM	Produced power forecast (Environmental)	In-plane solar radiation, cell temperature (Environmental) [24]
Aspatial	Hourly	SVM, FFANN, KNN, Deep ANN	Concentrated PV generation prediction (Environmental)	Concentrated solar radiation, non concentrated solar radiation, experiment time, PV module temperature (Environmental) [91]
Aspatial	Atemporal	FFANN, ANN	Load current of PV array, maximum power (Environmental, Specifications)	Load voltage, radiation level, temperature (Environmental) [60]
54 locations	Atemporal	FFANN	Estimation of sizing parameters (Specifications)	Altitude, longitude (Environmental) [52]
Aspatial	Hourly	Deep ANN	PV system real performance (Environmental)	Month, hour, angle, lab. performance, environment temperature, lab. temperature efficiency, local solar radiation (Environmental, Specifications) [4]
Aspatial	7.5 min	LSTM, RNN, FFANN, attention mechanisms	PV performance forecasting 1 step ahead (Environmental)	Temperature, Power time series (Environmental) [90]
Aspatial	Daily	FFANN	PV and Thermal system performance prediction and comparison: current, voltage, efficiency, thermal efficiency (Environmental)	Solar radiation, ambient temperature (Environmental) [3]
Aspatial	Atemporal	FFANN	Different PV systems output prediction (Environmental)	Light intensity, ambient temperature (Environmental) [82]
Single Point	Daily	FFANN	Maximum power generation of PV module (Specifications)	Irradiation, temperature, wind velocity, time parameter (Environmental) [29]
Single Point	Sub-hourly	FFANN	Maximum and normal operating power of PV module (Environmental, Specifications)	Solar radiation, module surface temperature (Environmental) [11]

Table 3.3: Tabla obtenida de Valderrama *et al.* [75], presenta propuestas de uso de ML para el problema de estimación de potencial técnico. En inglés del original. Descripción original: *Several ML application proposals for the technical potential estimation. The input and target variables classification is presented in parenthesis.*

El caso chileno

Las distintas propuestas presentadas en el Capítulo anterior nos dan cuenta de las diferentes avenidas posibles para abordar distintos aspectos del problema de la estimación del potencial de generación fotovoltaico. Claramente, muchas de estas avenidas serán restringidas por el tipo de datos, la disponibilidad que se tenga de estos y la resolución que estos tengan, tanto espacial como temporal. De la misma forma, el único estudio encontrado en literatura que realiza una estimación de potencial urbano chileno corresponde a Campos *et al.* [18], presentado en el Capítulo anterior. Como se describió previamente, el estudio realiza una estimación gruesa espacialmente (solo llegando a resolución a nivel de ciudad), utilizando fuertes supuestos de valores constantes a lo menos discutibles. Esta ausencia de estimación confiable y granular dificulta tanto la planificación de largo plazo, como la focalización de recursos y esfuerzos en zonas urbanas específicas con mejores potencialidades.

Recientemente, el Ministerio de Energía ha puesto a disposición una base de datos del potencial fotovoltaico urbano sobre Santiago, expresado a través del área favorable para instalaciones fotovoltaicas, de alta granularidad espacial, alcanzando una desagregación hasta manzanas individuales dentro de la ciudad. Por otro lado, el Servicio de Impuestos Internos (SII) dispone de una base de datos de propiedades que abarca prácticamente todo el territorio nacional, con una altísima desagregación espacial, llegando a nivel de rol tributario único. Esta base de datos contiene identificadores y direcciones que permiten, mediante procesamientos de tipo GIS, georreferenciar la base de datos, anclándola así a la realidad espacial. De igual interés, la base de datos contiene campos que permiten describir las edificaciones en los territorios concernidos: número de pisos de la edificación, superficie construida, materialidad, destino de la edificación, año de construcción o avalúo fiscal. Los registros resultan de particular interés, dado que varios pueden relacionarse directa o indirectamente con la forma tridimensional de las edificaciones, o con características de la distribución de las edificaciones dentro de una manzana, por lo cual pueden resultar útiles para el problema presente.

Mediante un proceso de georreferenciación de la base de datos SII, se ha logrado establecer un cruce espacial entre ambas bases de datos, con un nivel de identificación de manzana.

Es decir, se dispone de una base de datos consolidada sobre un área de Santiago, donde para cada manzana se conoce su potencial fotovoltaico, y una lista de edificaciones descritas en tal manzana por el SII. Esto permite vislumbrar un problema de aprendizaje, donde un modelo deba predecir el potencial de generación de una manzana a partir de la lista de atributos SII de esta, permitiendo a posteriori predecir sobre el resto del territorio utilizando la base de datos SII, alcanzando una resolución espacial de nivel de manzana, muchísimo más fino que lo actualmente disponible en el país. Sin embargo, igualmente surge inmediatamente una dificultad suplementaria: dado que distintas manzanas pueden contener un distinto número de edificaciones en su interior, nuestros datos serán de largo variable, lo cual dificulta el uso de algoritmos estándar de ML, al menos sin preprocesamientos previos.

Problema de aprendizaje general

Dado un fenómeno gobernado por una cierta distribución de probabilidad conjunta P sobre $\mathcal{X} \times \mathcal{Y}$ desconocida, entonces, una tarea de aprendizaje busca resolver el siguiente problema de minimización:

$$\min_{h:\mathcal{X}\rightarrow\mathcal{Y}} \int_{\mathcal{X}\times\mathcal{Y}} \ell(h(x), y) dP. \quad (5.1)$$

Donde $\ell : \mathcal{Y} \rightarrow \mathbb{R}_+^*$ es la llamada función de pérdida, que mide la diferencia entre las predicciones y los valores reales, usualmente obtenida a partir de alguna norma del espacio \mathcal{Y} .

Sin embargo, para siquiera intentar resolver este problema, se requeriría tener supuestos sobre la distribución P . En problemas de ML, solo podemos conocer P a través de muestras de datos obtenidos del fenómeno, usualmente denotadas (X, Y) , las cuales se separan habitualmente en conjuntos de entrenamiento (X_{tr}, Y_{tr}) , validación (X_{val}, Y_{val}) y test. Más aún, en la práctica no podemos resolver el problema de optimización sobre todas las posibles funciones, por lo cual previamente se restringe el espacio de búsqueda de la formulación general a un espacio de hipótesis $\mathcal{H} \subset \{\mathcal{X} \rightarrow \mathcal{Y}\}$. Así, los algoritmos de ML encuentran una hipótesis $h_{\mathcal{H}}^{tr}$ dados un espacio de hipótesis \mathcal{H} y un conjunto de entrenamiento como una solución del siguiente problema de optimización:

$$h_{\mathcal{H}}^{tr} := \arg \min_{h \in \mathcal{H}} \sum_{(x,y) \in (X_{tr}, Y_{tr})} \ell(h(x), y). \quad (5.2)$$

Esta optimización, sin embargo, dado un espacio de hipótesis lo suficientemente “grande” (es decir, con suficiente capacidad), siempre entregará una función que se ajusta perfectamente a los datos de ejemplo, pero de mucho peor desempeño en datos futuros: estamos en la presencia del fenómeno de sobreajuste. Dado que no queremos aprender de “memoria” los datos presentes, si no predecir datos futuros, el algoritmo de optimización busca la hipótesis en función de la pérdida sobre el conjunto de entrenamiento, pero utilizamos el conjunto de

validación para evaluar el desempeño relativo de distintas funciones de hipótesis, algoritmos de optimización y espacios de hipótesis. El conjunto de test o pruebas es reservado para realizar una evaluación final de los modelos una vez ajustados y entrenados, mientras el algoritmo elige la hipótesis a partir del problema de optimización anterior restringido al conjunto de entrenamiento. En caso de utilizar el conjunto de validación, este será utilizado por ende para la optimización de los hiperparámetros del algoritmo o la selección de distintas familias de posibles modelos.

En nuestro caso, buscaremos la mejor representación interna de la red para el problema en particular. Modificar la representación de la red equivale a modificar el espacio de hipótesis, por lo que si denotamos por

$$\Xi_r^{\mathcal{H}} := \{h \in \mathcal{H} : h \text{ usa representación } r\}. \quad (5.3)$$

Entonces queremos resolver:

$$\min_r \sum_{(x,y) \in X_{val}^r, Y_{val}} \ell \left(h_{\Xi_r^{\mathcal{H}}}^{tr}(x), y \right), \quad (5.4)$$

donde r toma valores “homogéneo”, “heterogéneo” o “mixto”, los cuales definiremos más adelante.

Convoluciones 1D

Las redes neuronales convolucionales, originalmente inspiradas por comportamientos observados en la corteza visual de animales [33], tienen una arquitectura bien condicionada para el reconocimiento de patrones en imágenes en el caso dos dimensional, o en secuencias en el caso uno dimensional. Esto lo logran gracias al uso de la operación de convolución, la cual aplica una combinación fija de pesos entrenables (o kernel usualmente) sobre las distintas posiciones en una secuencia. Esto permite por ende obtener el mismo valor en la secuencia de salida frente al mismo patrón en los datos de entrada, independiente de la posición del patrón dentro de la secuencia. En el caso de secuencias de entrada multidimensionales, el kernel tiene distintos grupos de pesos asociados a cada dimensión de la entrada.

La principal operación de las redes convolucionales es la convolución. En el caso general, esta opera sobre una secuencia de entrada arbitraria $\xi : \{1, \dots, \alpha\} \rightarrow \mathbb{R}^\beta$ de largo $\alpha \in \mathbb{N}$ y $\beta \in \mathbb{N}$ dimensiones. La convolución viene definida por un kernel denotado $g : \{1, \dots, \gamma\} \rightarrow \mathbb{R}^\beta$, de tamaño $\gamma \in \mathbb{N}$. Notemos que su dimensionalidad corresponde a la dimensionalidad de la secuencia de entrada, y su largo γ es fijo y debe ser menor o igual al largo α de la secuencia de entrada. Entonces, la convolución $\xi * g$ es definida mediante la siguiente fórmula:

$$(\xi * g)(i) = \sum_{k=1}^{\beta} \left(\sum_{p=1}^{\alpha} \xi^{(k)}(p) g^{(k)}(i - p + 1) \right), \quad (6.1)$$

donde denotamos por $g^{(i)}(j)$ la j -ésima posición de la secuencia g en la i -ésima dimensión, $i \in \{1, \dots, \alpha\}$, y tomamos la convención de $g(i) = 0$ para $i \notin \{1, \dots, \gamma\}$.

Notemos que si bien la suma interior itera sobre α índices distintos, en la práctica solo γ valores de ξ contribuyen al resultado para cada posición, dado el dominio de definición de g .

Ahora, habiendo definido la operación convolución, podemos definir la secuencia de salida de un filtro convolucional, o neurona en el contexto de las redes convolucionales. Si el filtro viene definido por el kernel g , el que opera sobre el input ξ , entonces la salida y puede ser computada siguiendo la fórmula:

$$y(i) = \sigma(b + (\xi * g)(i)), \quad (6.2)$$

donde $\sigma : \mathbb{R} \rightarrow \mathbb{R}$ es una función de activación, habitualmente no lineal, y $b \in \mathbb{R}$ una constante.

Para definir una capa convolucional, distintos filtros, o neuronas, son aplicados paralelamente sobre la secuencia de entrada, obteniendo como salida una secuencia multidimensional, con dimensión igual a la cantidad de neuronas. Esto permite extraer simultáneamente distintos patrones de las secuencias de entrada, los cuales posterior aplicación de la función de activación, componen los llamados mapas de atributos. Más formalmente, la salida Y de una capa neuronal con n neuronas, donde $Y^{(j)}$ denota la j -ésima dimensión de salida y la j -ésima neurona viene definida por el kernel $g^{(j)}$ y el intercepto $b^{(j)}$, puede ser computada mediante la siguiente fórmula:

$$Y^{(j)}(i) = \sigma(b^{(j)} + (\xi * g^{(j)})(i)). \quad (6.3)$$

Recordemos que aquí, tanto los pesos de los filtros convolucionales como los interceptos son valores entrenables, mientras que la función de activación, el largo del filtro convolucional y el número de neuronas son hiperparámetros a determinar previo al entrenamiento.

Finalmente, una red convolucional es definida mediante una sucesión de capas convolucionales, donde la entrada de una capa corresponde a la salida de la capa anterior. Esta arquitectura ha demostrado en literatura permitir a la red aprender características significativas de los datos [42] de manera secuencial. Así, la red al aprender obtiene representaciones convolucionales internas de los datos de un nivel de abstracción creciente.

Propuesta e hipótesis

En este capítulo se describe detalladamente la propuesta del presente trabajo. Para esto, definiremos primeramente las distintas representaciones de los datos disponibles que se busca estudiar, posteriormente esbozaremos hipótesis de investigación y finalmente presentaremos los objetivos para decidir la hipótesis.

En el caso presente, los N datos corresponden a pares (x_n, y_n) , donde x_n viene dado por los descriptores presentes en la base de datos SII (número de pisos, superficie construida y avalúo fiscal) del conjunto de edificaciones asociadas a la n -ésima manzana, mientras y_n corresponde al área favorable para instalaciones fotovoltaicas en la manzana, la cual puede fácilmente ser transformada a un potencial de generación utilizando las bases de datos de potencial teórico y herramientas de estimación técnica del Ministerio de Energía. Si definimos como μ_n el número de edificaciones presentes en la n -ésima manzana, y como $J \in \mathbb{N}$ el número de características descriptoras de los edificios, tendremos que, para cada n ,

$$x_n \in \mathbb{R}^{\mu_n \times J}. \quad (7.1)$$

En otras palabras, x_n está compuesto de μ_n edificaciones, cada una con J atributos descriptores. Por otro lado, $Y \in \mathbb{R}^N$, dado que cada $y_n \in \mathbb{R}$. Optamos por realizar *zero-padding* sobre nuestros datos de entrada, es decir agregamos valores cero a los registros para obtener un largo uniforme. Si definimos $L = \max_n \{\mu_n\}$, entonces nuestra representación secuencial de largo uniforme \bar{x}_n viene definida mediante la siguiente fórmula:

$$\bar{x}_n^{(j)}(i) = \begin{cases} x_n(i, j) & \text{si } 1 \leq j \leq \mu_n \\ 0 & \text{e.o.c} \end{cases}, \quad (7.2)$$

donde $i \in \{1, \dots, L\}$ e $i \in \{1, \dots, J\}$. Así, tenemos que $\{\bar{x}_n\}_n \subset \mathbb{R}^{L \times J}$. Consideramos cada uno de estos ejemplos como una secuencia, donde el índice correspondiente a los distintos edificios dentro de cada manzana corresponde al índice de la secuencia, y el segundo índice asociado a los descriptores corresponde a las distintas dimensiones de esta secuencia.

Definiremos la representación de los datos por \bar{x}_n como la representación de tipo homogénea, dado que las distintas posiciones de la secuencia representando cada manzana accede a distintas instancias de la misma entidad, es decir, al avanzar por la secuencia estamos accediendo al mismo *tipo* de elementos, en este caso a distintas edificaciones. Definamos igualmente la siguiente representación:

$$\tilde{x}_n^{(i)}(j) = \bar{x}_n^{(j)}(i), \quad (7.3)$$

donde notamos que los roles jugados por el índice de posición y el índice de dimensión son intercambiados. Esta nueva representación $\tilde{x}_n \in \mathbb{R}^{J \times L}$ la denotaremos como la representación heterogénea de datos. La denotamos como heterogénea, dado que distintas posiciones en la secuencia nos entregan distintos *tipos* de datos, mientras las distintas dimensiones acceden a las distintas instancias de edificaciones del ejemplo particular.

Al reescribir las operaciones convolución descritas en la sección anterior para ambos casos, obtenemos la operación convolución sobre la secuencia homogénea y heterogénea respectivamente, como se presenta a continuación.

$$\begin{aligned} \text{Homogénea: } (\bar{x}_n * g)(i) &= \sum_{k=1}^J \left(\sum_{p=1}^L \bar{x}_n^{(k)}(p) g^{(k)}(i - p + 1) \right), \\ \text{Heterogénea: } (\tilde{x}_n * g)(i) &= \sum_{k=1}^L \left(\sum_{p=1}^J \tilde{x}_n^{(k)}(p) g^{(k)}(i - p + 1) \right). \end{aligned} \quad (7.4)$$

Para notar las peculiaridades de cada representación, expresemos estas ecuaciones en el caso particular donde el largo del kernel g es igual a 1. Dada la definición, notamos que la suma interior evalúa 0 para todos los valores donde $i - p + 1 \neq 1$, es decir, esta se reduce al elemento $p = i$ y podemos reescribir las operaciones como:

$$\begin{aligned} \text{Homogénea: } (\bar{x}_n * g)(i) &= \sum_{k=1}^J (\bar{x}_n^{(k)}(i) g^{(k)}(1)), \\ \text{Heterogénea: } (\tilde{x}_n * g)(i) &= \sum_{k=1}^L (\tilde{x}_n^{(k)}(i) g^{(k)}(1)) \\ &= \sum_{k=1}^L (\bar{x}_n^{(i)}(k) g^{(k)}(1)). \end{aligned} \quad (7.5)$$

De estas ecuaciones se puede notar la principal diferencia en términos de la operación convolución de ambas representaciones. En el caso de la representación homogénea, para una posición i estamos combinando los J descriptores asociados con la edificación i -ésima, utilizando los mismos valores del kernel g para las distintas edificaciones a medida iteramos sobre i . En cambio, en la representación heterogénea, al calcular una posición i estamos combinando mediante el kernel g todas las edificaciones presentes, utilizando solamente el i -ésimo descriptor presente. Igualmente, distintas posiciones i en el caso heterogéneo acceden a

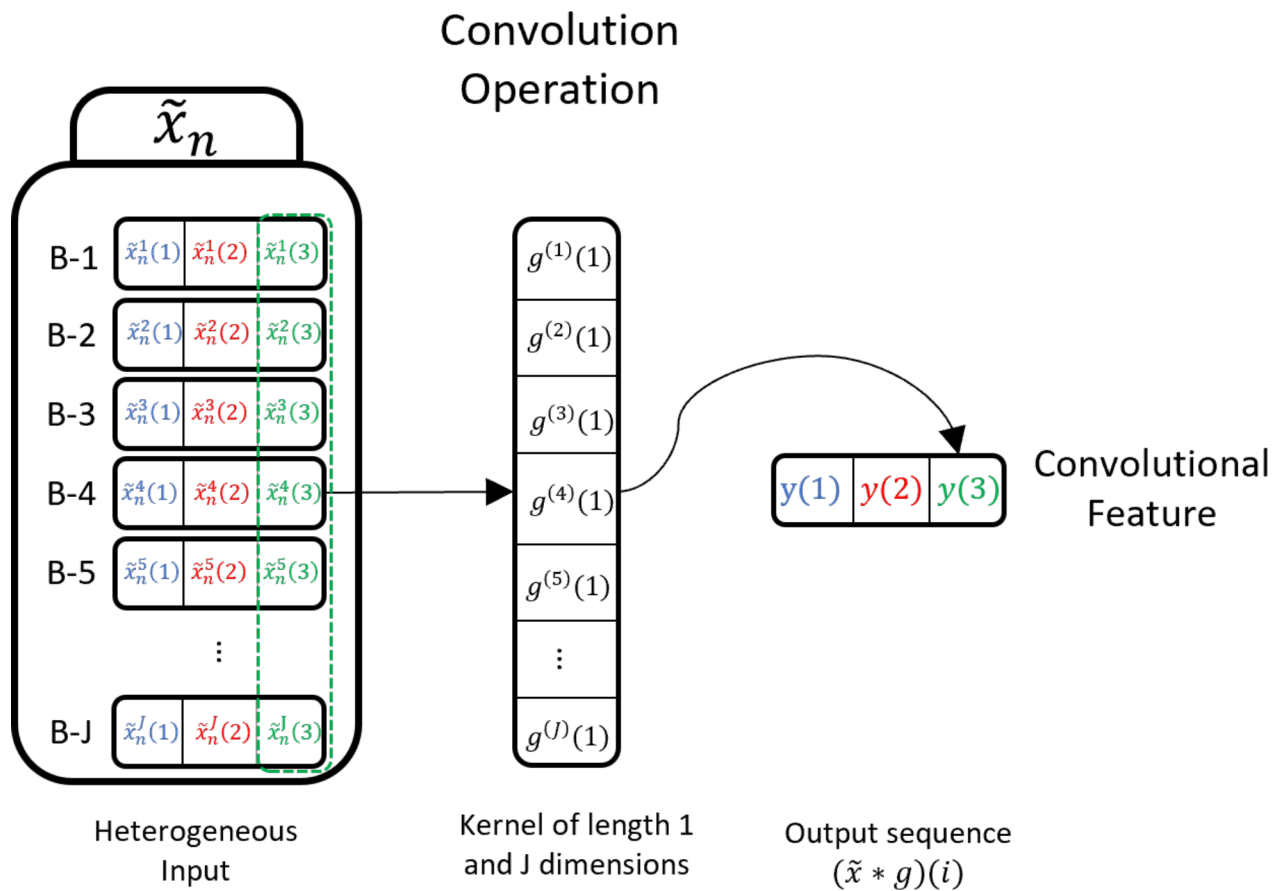


Figura 7.1: Representación esquemática de la operación convolucional de una sola neurona sobre la representación heterogénea de una manzana particular. La definición matemática de la convolución presentada se encuentra descrita en la ecuación (7.5). Obtenida de Valderrama *et al.* [76], en inglés del original.

distintos descriptores de todos los edificios simultáneamente. Podemos ver una representación gráfica de los datos y de la operación convolución heterogénea en la Figura 7.1.

Podemos notar que la representación heterogénea utiliza los mismos pesos convolucionales para distintos tipos de datos, sin embargo, opera simultáneamente sobre todas las edificaciones, mientras que la homogénea utiliza cada peso de la convolución sobre el mismo tipo de dato, pero secuencialmente, edificio por edificio. Se podría suponer que la representación homogénea tradicional facilita a la red la descripción resumida de edificios individuales a partir de combinaciones significativas de sus descriptores, mientras la representación heterogénea fuerza a la red a encontrar pesos que puedan resumir características de toda la manzana al combinar todos los edificios simultáneamente. Adicionalmente, en los casos donde la cantidad de dimensiones de la representación original sea mucho menor que el largo de las secuencias, como es el caso presente, la representación heterogénea reducirá la dimensionalidad de la representación interna de la red al compararla con la representación homogénea, dado que, como presenta la figura 7.1, el largo de la representación obtenida por cada neurona convolucional será igual al número de dimensiones de la representación original.

Se propone un tercer tipo de red, combinando ambas representaciones. Esta red, denominada mixta, utilizará en sus capas iniciales representaciones homogéneas que permitan aprender representaciones de los edificios individuales de la manzana, seguido de una transposición de las representaciones para continuar las capas más profundas utilizando una representación heterogénea que permita resumir toda la manzana combinando todas las edificaciones

simultáneamente.

Proponemos explorar este tipo de representaciones internas de la red, tanto homogéneas, heterogéneas y mixtas de manera sistemática. En específico, se propone, comparar sistemáticamente representaciones heterogéneas, homogéneas y mixtas, para evaluar la de mejor desempeño para el problema particular de la estimación fotovoltaica urbana a nivel de manzana utilizando datos de las edificaciones. Esta evaluación se realizará tanto en términos de la precisión alcanzada por las distintas arquitecturas, como del costo computacional asociado a estas. Adicionalmente, los experimentos a realizar permiten comparar las distintas representaciones internas de los datos aprendidas por la red.

La principal hipótesis que se busca decidir a partir de la presente propuesta es:

Las redes convolucionales uno dimensionales mixtas obtienen mejores desempeños en el problema de estimación de generación fotovoltaica urbana sobre los techos de Santiago al compararlas con representaciones heterogéneas y el estado del arte homogéneo, en términos del error absoluto porcentual medio.

Adicionalmente, en función de los hallazgos realizados durante la experimentación, se propone evaluar una hipótesis secundaria, presentada a continuación:

Para el problema de estimación de generación fotovoltaica urbana sobre los techos de Santiago, la representación heterogénea propuesta es más eficiente computacionalmente que la representación homogénea del estado del arte, sin pérdidas significativas de precisión.

Para lograr decidir estas hipótesis, se proponen los siguientes objetivos específicos.

1. Encontrar configuraciones óptimas de arquitecturas convolucionales uno dimensionales tanto mixtas, heterogéneas y homogéneas, para el problema en cuestión, mediante una comparación sistemática de los desempeños sobre el conjunto de validación.
2. Entrenar redes con desempeños óptimos para cada representación en el problema en cuestión, aprendiendo transformaciones de los datos de entradas a representaciones convolucionales de largo fijo.
3. Comparar las ventajas comparativas de las distintas representaciones mixtas, homogéneas y heterogéneas mediante la comparación de desempeños de arquitecturas con sus hiperparámetros optimizados, tanto en términos de los errores como en términos de costos computacionales.
4. Evaluar la capacidad de las representaciones convolucionales de facilitar el aprendizaje del problema en cuestión a algoritmos mal condicionados para datos de esta índole a partir del desempeño de regresiones lineales sobre las representaciones internas.

Experimentos

En el presente capítulo se describe el trabajo experimental realizado. En particular, se describe inicialmente el conjunto de datos disponible para experimentación, seguido de la presentación del proceso de optimización de hiperparámetros de cada tipo de representación utilizada. Posteriormente se presenta una comparación del desempeño y características de cada una de las redes optimizadas, seguido de una comparación de los costos computacionales de las distintas aproximaciones, para finalmente presentar un análisis del desempeño de las representaciones internas aprendidas por las redes.

8.1. Conjunto de datos

El conjunto de datos experimentales se compone de 5018 ejemplos, cada uno de estos corresponde a una manzana de ciudad de Santiago. Cada uno de los ejemplos consta de una estimación del área favorable de la manzana, proporcionada por el Ministerio de Energía, y una lista de atributos de los edificios de la manzana asociados a registros del Servicio de Impuestos Internos. Estos atributos corresponden a la suma de los avalúos fiscales de las propiedades asociadas a cada edificio, suma de las superficies construidas asociadas a las propiedades asociadas a cada edificación y el número de pisos de cada edificio.

Este conjunto de datos es separado en conjuntos de entrenamiento, validación y pruebas siguiendo una proporción de 75 %, 15 % y 10 % respectivamente, obteniendo un total de 3763 ejemplos de entrenamiento, 753 ejemplos de validación y 502 ejemplos de pruebas.

8.2. Ajuste de hiperparámetros

Primeramente, para la exploración y comparación de desempeño de redes convolucionales mixtas frente a sus pares heterogéneos u homogéneos, se propone realizar una exploración sistemática de los hiperparámetros de estos tres tipos de redes convolucionales bajo una es-

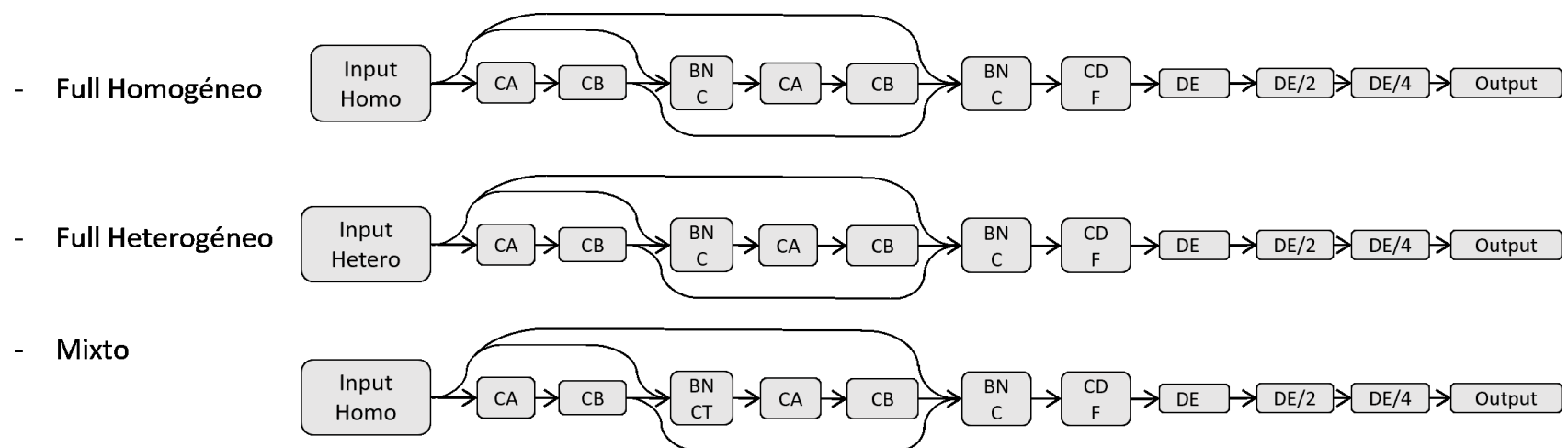


Figura 8.1: Presenta una representación esquemática de las tres arquitecturas sobre las que se realizó optimización de los hiperparámetros. Cada casilla comenzando con una C corresponde a una convolución, BN corresponde a una capa de *Batch Normalization*, las flechas curvas corresponden a *skip connections*, la letra T denota a una transposición, y D a capas densas. Las letras A, B, C y E corresponden a los hiperparámetros a optimizar, y representan el número de neuronas de la capa en cuestión.

Representación	Val Loss
Homogénea	$0,45 \pm 0,01$
Heterogénea	$0,47 \pm 0,03$
Mixta	$0,48 \pm 0,011$

Tabla 8.1: Resultados de exploración de representaciones sobre validación utilizando 5-fold.

estructura previamente definida. Esta estructura se presenta en la figura 8.1 donde cada casilla comenzando con una C corresponde a una convolución, BN corresponde a una capa de *Batch Normalization* [35], las flechas curvas corresponden a *skip connections* [28], la letra T corresponde a una transposición y D a capas densas. Esta arquitectura de red para el problema específico en cuestión fue obtenida en el contexto de la publicación en conferencia por Valderrama *et al.* [76] asociada igualmente a esta iniciativa. Por otro lado, las letras A, B, C y E corresponden al número de neuronas de cada capa, y corresponden a los hiperparámetros a optimizar.

Se exploran estos distintos hiperparámetros, tomando potencias de 2 con exponente variando entre 2 y 8, dando un total de combinaciones a explorar de 2744 para cada tipo de red. Igualmente, se utiliza *5-folds* sobre el conjunto de entrenamiento y validación al momento de realizar la búsqueda grillada sobre los hiperparámetros. Esta búsqueda se realiza independientemente para cada una de las representaciones, obteniendo por ende una configuración óptima en términos de los hiperparámetros para cada tipo de representación.

Los resultados en término del desempeño de la mejor configuración promediado sobre los *5-folds* realizados se presenta en la tabla 8.1. De esta tabla podemos ver que los resultados en términos de la pérdida son relativamente cercanos para todas las redes. De todas formas, siendo estos resultados obtenidos sobre el mismo conjunto de validación con que se optimizan las redes, no se puede concluir respecto de la capacidad de generalización inherente de cada una. Por otro lado, en la tabla 8.2 se puede apreciar los valores obtenidos para cada uno de

Hiperparámetro	Homogénea	Heterogénea	Mixta
A	128	128	16
B	16	128	32
D	32	256	256
E	64	32	16

Tabla 8.2: Resultados de hiperparámetros óptimos.

los hiperparámetros de la configuración óptima de cada tipo de red.

8.3. Comparación redes optimizadas

Habiendo optimizado la arquitectura para cada uno de los tipos de representación de los datos, se comparan las mejores de cada categoría sobre el conjunto de prueba. Para esto, se entrena la red utilizando el conjunto de entrenamiento, evaluando en cada época de entrenamiento el error sobre el conjunto de validación, y deteniendo el entrenamiento luego de 10 épocas consecutivas sin presentar mejoras en el error sobre el conjunto de validación. Una vez entrenada la red, se recuperan los pesos obtenidos para la época con menor error de validación, y con estos se calculan las métricas de error sobre el conjunto de pruebas. Este proceso se realiza para las 3 representaciones utilizando los hiperparámetros optimizados presentados en la tabla 8.2, y se realizan 10 repeticiones inicializando aleatoriamente los pesos iniciales de las redes. Para el efecto de la comparación se utilizará la métrica del error porcentual medio o *Mean Absolute Percentage Error* (MAPE).

Los resultados de estos experimentos, en términos de la pérdida y el MAPE, sobre el conjunto de test, son presentados en la tabla 8.3. De la tabla podemos apreciar primeramente, que la representación que obtiene los mejores desempeños, tanto en términos de la pérdida como del MAPE es la representación mixta. Cabe notar, sin embargo, que las diferencias entre los desempeños son pequeñas en magnitud.

De hecho, al realizar los test de significancia estadística sobre los resultados de errores presentados en la tabla 8.3, para las 10 repeticiones y una significancia del 5%, utilizando como hipótesis nula que la media de las distribuciones de los distintos errores es igual, obtenemos los siguientes resultados. No logramos rechazar la hipótesis nula para la comparación entre la red homogénea y mixta ni para la comparación entre las redes homogénea y heterogénea, es decir, no se logra apreciar una diferencia estadística entre la performance de las redes del estado del arte (homogénea) y las propuestas en este contexto (heterogéneas y mixtas). Sin embargo, al comparar la red mixta y la heterogénea, la hipótesis nula puede rechazarse y obtenemos que la red con mejor desempeño es la red mixta. Los *p-values* asociados a los test estadísticos son los siguientes: 0.1326 para el par homogénea y mixta, 0.0710 para el par homogénea y heterogénea y finalmente menos de 0.00001 para el par mixta y heterogénea.

Por otro lado, de la tabla podemos apreciar igualmente que el número de dimensiones de la representación interna de la red presentan una variabilidad muy grande en función de la representación utilizada. Notamos primeramente que la red mixta tiene una representación interna aproximadamente la mitad de grande que la red homogénea, mientras la red hetero-

Representación	NDim	Test Loss	Test MAPE
Homogénea	63296	0,7864 \pm 0,0071	11,53 % \pm 0,028 %
Heterogénea	192	0,8021 \pm 0,0075	11,56 % \pm 0,022 %
Mixta	32869	0,7791 \pm 0,0024	11,51 % \pm 0,009 %

Tabla 8.3: Resultados de redes optimizadas sobre test, en términos de la pérdida y el MAPE. Adicionalmente se presenta la dimensionalidad NDim de la representación interna de los datos en la última capa previo a la capa de *flatten* de cada arquitectura. Los intervalos de confianza consideran un 5 % de incertidumbre.

génea obtiene una representación interna con menos del 1 % de las dimensiones de las otras dos representaciones.

Esto se debe a que la operación convolucional sobre la representación homogénea no reduce significativamente la dimensionalidad de las secuencias, dado que las secuencias homogéneas tienen mayor largo que número de dimensiones. De manera complementaria, las convoluciones heterogéneas en este caso mantienen una dimensionalidad mucho menor dado que las secuencias heterogéneas de entrada tienen tan solo largo 3 y dimensionalidad mucho más alta, la cual es rápidamente resumida por la operación convolución. Finalmente, en el caso mixto, las dimensiones altas vuelven a aparecer en la representación de salida dado que las *skip connections* presentes en la red las preservan independiente de las modificaciones dimensionales en secciones posteriores de la red.

Resulta interesante notar que, si bien la red heterogénea es aquella con peores desempeños, alcanza desempeños en el mismo orden de magnitud que las otras representaciones, sin si quiera poder diferenciarlos estadísticamente de los desempeños del estado del arte (red homogénea), mientras utiliza una representación interna órdenes de magnitud menos dimensional. En términos porcentuales, la red heterogénea obtiene sus resultados utilizando menos de un 0,4 % de las dimensiones que utiliza la red homogénea y menos de un 0,6 % de las dimensiones de la representación interna de la red mixta. Naturalmente, surge la pregunta de los costos computacionales asociados a representaciones de tan distinta dimensionalidad.

8.4. Comparación costos computacionales

De la sección anterior se nota que si bien los resultados de las arquitecturas propuestas en el presente trabajo no logran superar los desempeños de la representación homogénea habitualmente utilizada, si presentan una diferencia radical en la dimensionalidad de la representación. En particular, se aprecia que la red heterogénea obtiene desempeños equivalentes estadísticamente al estado del arte homogéneo, utilizando una representación órdenes de magnitud menor. A continuación, presentamos resultados de la evaluación de los costos computacionales asociados a las arquitecturas propuestas. Los experimentos referidos en esta sección fueron ejecutados en una máquina Apple, Macbook pro, procesador M1 pro con 16 GB de RAM.

Para evaluar los costos computacionales, evaluamos primeramente los tiempos de entrenamiento de las distintas arquitecturas optimizadas para cada una de las representaciones. Para

Representación	Costo Entrenamiento (s)	Costo predicción (s)
Homogénea	30,6301 ± 3,8750	1,3644 ± 0,0356
Heterogénea	12,4897 ± 0,7603	0,3181 ± 0,0095
Mix	73,9660 ± 18,8049	1,4682 ± 0,0573

Tabla 8.4: Presenta los tiempos promedio de entrenamiento y predicción de las distintas arquitecturas, en segundos, con intervalos de confianza sobre las 10 repeticiones con una significancia del 5 %.

Representación	NDim	Test MSE	Test MAPE
Homogénea	63296	1159 ± 136,9	1070,19 % ± 137,44 %
Heterogénea	192	6,06 ± 0,114	13,52 % ± 0,11 %
Mixta	32869	1775 ± 154,88	1484,01 % ± 118,12 %

Tabla 8.5: Resultados de uso de representaciones intermedias de la red para ajuste de regresión lineal, con intervalos de confianza sobre las 10 repeticiones con una significancia del 5 %.

esto, se entrena una red sobre los 3.763 ejemplos de entrenamiento, utilizando *EarlyStopping* con paciencia igual a 5 épocas, sobre el conjunto de 753 ejemplos de validación, evaluando el tiempo de entrenamiento. Esto se repite 10 veces, con inicialización aleatoria de la red, y se registra el tiempo de entrenamiento para cada una de las 3 distintas arquitecturas.

De igual manera, para cada una de las redes entrenadas, se evalúa el tiempo de predicción. Para esto, se registra el tiempo que toma la red en predecir el valor de los conjuntos de entrenamiento y validación, es decir el tiempo de evaluación de un total de 4516 ejemplos.

Los resultados de estas evaluaciones son presentados en la Tabla 8.4. La tabla presenta los valores promedio sobre las 10 repeticiones de los tiempos de computo para entrenamiento y predicción para cada una de las arquitecturas, en segundos. Se presentan intervalos de confianza con una significancia del 5 %. Se aprecia claramente que la red heterogénea es la más eficiente en términos computacionales, tanto en el entrenamiento como en la predicción de la red ya entrenada.

En efecto, se realiza el test estadístico con la hipótesis nula de medias de la distribución de los tiempos iguales, con significancia del 5 %. La hipótesis nula se rechaza cuando comparamos la heterogénea con la mixta y la heterogénea con la homogénea. Adicionalmente, la comparación entre la red mixta y la homogénea igualmente rechaza la hipótesis nula. Es decir, la representación homogénea es menos costosa computacionalmente, tanto en entrenamiento como predicción, que la representación mixta. Sin embargo, la representación heterogénea es más eficiente que la propuesta mixta y que el estado del arte homogéneo, tanto para el entrenamiento como para la predicción, siendo así la más eficiente de las 3 representaciones.

8.5. Representaciones internas

Para entender mejor el desempeño relativo de las representaciones internas de las redes, se ejecutó una segunda batería de experimentos. Para estos, se utilizaron las representaciones intermedias de las redes optimizadas, completamente entrenadas. Se utilizan los valores

internos de la red (representaciones convolucionales) como codificaciones de largo fijo de los datos de entrada.

Sobre estas representaciones aprendidas se ajustan regresiones lineales, las que permiten de manera rápida evaluar la capacidad de describir el fenómeno en cuestión que cada representación del problema es capaz de alcanzar. Se utiliza como representación aprendida de cada red la última capa previo a la capa de *flatten* de esta, representada como un vector. Igualmente, se generará una regresión lineal por cada una de las 10 repeticiones del entrenamiento de las redes, para obtener intervalos de confianza equivalentes a los presentados en la sección previa.

Los resultados, en términos del error cuadrático medio y el MAPE sobre los conjuntos de test, para las regresiones lineales ajustadas utilizando las representaciones internas de cada una de las redes, se presentan en la tabla 8.5, junto con el número de dimensiones “NDim” de la representación en cuestión.

Se nota que los errores obtenidos por la representación homogénea y heterogénea son órdenes de magnitud mayores que la representación heterogénea y que las redes optimizadas. Esto puede darse en gran parte por no estar bien condicionada la regresión lineal sobre representaciones altamente dimensionales, pero igualmente podríamos hipotetizar que la sección densa juega un rol más importante que la sección convolucional en las redes mixtas y homogéneas. Independiente, resulta nuevamente destacable el desempeño de la representación heterogénea. En el caso de la regresión lineal con representación heterogénea, obtenemos el mejor resultado de las representaciones aprendidas, obteniendo un MAPE tan solo un par de puntos porcentuales por sobre el desempeño de las redes presentadas en la sección previa.

Es decir, la red heterogénea encuentra una representación con órdenes de magnitud menos dimensiones que las otras redes, y que permite llevar el problema a una solución lineal con una pérdida de desempeño de tan solo un par de punto porcentuales. Podemos hipotetizar que la red heterogénea logra aprender una representación interna más eficiente que las otras redes.

Contribuciones

El trabajo presentado en el presente documento, ha sido publicado en medios académicos. Primeramente, una publicación técnica de conferencia presentó los resultados de la experimentación preliminar:

- Alvaro Valderrama, Carlos Valle, Marcelo Ibarra, y Héctor Allende. A heterogeneous 1d convolutional architecture for urban photovoltaic estimation. Conferencia Intelligent Computing Theories and Application 2021. [76]

Luego, una revisión exhaustiva del estado del arte y sus características permitió la publicación de un *survey* de las distintas aplicaciones de ML para la estimación del potencial fotovoltaico, en la revista *Neurocomputing* indexada por *Web of Science* (WOS):

- Alvaro Valderrama, Carlos Valle, Hector Allende, Marcelo Ibarra, and Camila Vásquez. Machine learning applications for urban photovoltaic potential estimation: A survey. Revista Neurocomputing. 2023. [75]

De los resultados de los experimentos se pueden obtener ciertas conclusiones y emitir ciertos juicios respecto de las hipótesis planteadas en el Capítulo 7. A continuación presentamos estos hallazgos y discutimos su significancia y futuras avenidas de investigación que se perfilan de estas.

Primeramente, se nota que al evaluar la significancia estadística de los desempeños comparativos entre las arquitecturas mixtas y el estado del arte homogénea, no se logra apreciar una diferencia. Es decir, no se puede afirmar a partir de los resultados de la presente investigación que la representación mixta sea mejor que la representación homogénea. Por ende, la primera hipótesis se rechaza, dado que no se logra demostrar diferencias significativas en términos del error entre ambas representaciones. Al comparar la representación heterogénea con la estándar homogénea, tampoco apreciamos diferencia significativa, por lo cual ninguna de las dos representaciones innovadoras propuestas logra mejorar los desempeños en términos del MAPE.

Esto es distinto a lo evaluado en la publicación de conferencia [76], la que inspira la actual iniciativa, en la cual se evidencia un mejor desempeño de la representación heterogénea frente a la representación homogénea. Esta diferencia puede deberse a dos diferencias fundamentales de los experimentos realizados. Primeramente, en la publicación referenciada, solo se disponía de una base de datos preliminar con 196 ejemplos en total, todos correspondientes a un área acotada y relativamente homogénea de Santiago. La propuesta actual utiliza 5018 ejemplos repartidos en distintas comunas de Santiago, dando por un lado mayor robustez al aprendizaje, y mayor variabilidad entre los datos disponibles. Por otro lado, dada la cantidad reducida de datos de entrenamiento, en el contexto de la publicación previa, no se pudo realizar una optimización de los hiperparámetros de las distintas redes, utilizando una configuración diseñada manualmente, tanto para el caso homogéneo como heterogéneo. En cambio, en la presente iniciativa, esta optimización de hiperparámetros se realizó, y las comparaciones se realizan sobre las distintas configuraciones utilizando para cada una de ellas su configuración óptima en términos de hiperparámetros.

Si bien no podemos por ende afirmar que las representaciones novedosas tienen mejores desempeños que el estado del arte, tampoco tienen peores desempeños. Esto se vuelve relevante al evaluar el número de dimensiones de la representación interna de la red: en el caso mixto esta tiene aproximadamente la mitad de dimensiones que la red homogénea, mientras la red heterogénea utiliza menos de un 1 % de las dimensiones que utiliza la red homogénea. Una diferencia tan significativa resulta interesante dado que los resultados de las distintas representaciones son equivalentes, utilizando algunas redes representaciones internas mucho más compactas.

Este hecho se ve igualmente reflejado en los costos computacionales de la red heterogénea. Al comparar los tiempos de entrenamiento y de predicción de las tres representaciones, la red heterogénea tiene los mejores costos computacionales por lejos. La diferencia entre la red heterogénea y homogénea es estadísticamente significativa, por lo que no se rechaza la segunda hipótesis propuesta en la sección 7. Igualmente curioso resulta el costo de entrenamiento de la red mixta, el cual es superior a cualquiera de las otras dos aproximaciones, a pesar de utilizar representaciones internas de menor dimensión que la red homogénea. Independiente de esto, se logra demostrar que una red utilizando una representación no habitual en el estado del arte, logra los mismos desempeños en términos de error (MAPE) que el estado del arte, con costos computacionales significativamente mejores.

Más aún, la eficiencia de la representación heterogénea se evidencia nuevamente al evaluar el desempeño de las representaciones aprendidas. Al evaluar estas representaciones aprendidas en términos de su desempeño predictivo mediante el uso de una regresión lineal, observamos que la representación heterogénea, a pesar de tener órdenes de magnitud menos dimensiones que las otras representaciones aprendidas, logra alcanzar el mejor desempeño en este esquema por un amplio margen. Más aún, la regresión lineal sobre la representación aprendida heterogénea obtiene desempeños tan solo 2 puntos porcentuales peores que las redes convolucionales, lo cual resulta destacable en el contexto de un problema altamente no lineal. Es decir, en el caso de la red heterogénea podemos hipotetizar que gran parte del aprendizaje del fenómeno ocurre en las capas convolucionales, obteniendo una representación resumida y razonablemente lineal del problema.

Trabajo futuro

Los hallazgos descritos previamente abren interesantes avenidas de investigación e incógnitas, algunas de las cuales se presentan a continuación.

Primeramente, surge la incógnita de por qué la representación mixta, si bien utiliza representaciones internas de menor dimensionalidad que la representación homogénea, es más costosa de entrenar. Esto podría deberse a algún condicionamiento del flujo de los gradientes dada la transposición intermedia en la red, o simplemente a particularidades de la implementación del gradiente descendiente para este tipo de operaciones.

Igualmente interesante resultaría evaluar en qué otro tipo de datos la red heterogénea puede lograr desempeños equivalentes al estado del arte. En el caso particular de este problema, la representación secuencial de los edificios tiene la característica de no poder asignar un orden natural *a priori*, dado que no tenemos información de la distribución de las edificaciones al interior de la manzana. Surge la duda si el comportamiento de la red heterogénea frente a las homogéneas viene dado por esta falta de orden inducido en los datos o si es observable para series ordenadas. La pregunta natural que surge es como se comporta la red heterogénea en problemas de series de tiempo, donde el orden de la serie es intrínseco.

Finalmente, otra avenida de investigación corresponde a encontrar otras representaciones de tipo mixta que si permitan mejorar los desempeños de la red, tanto para este problema en particular como para otros problemas que admitan representaciones secuenciales. Se pueden explorar otras arquitecturas que permitan combinar las representaciones heterogéneas y homogéneas, con dos secciones iniciales “paralelas” que procesen los input independientemente. Esto permitiría a cada sección aprender una representación útil heterogénea u homogénea con menor dependencia de la otra representación. Posteriormente ambas representaciones se pueden combinar mediante una sección densa que permita realizar la predicción. Otra idea que se podría evaluar es la de pre-entrenar redes homogéneas y heterogéneas similares a las propuestas en este proyecto, para posteriormente utilizar las secciones convolucionales pre-entrenadas en una red mixta, que combine ambas representaciones, pudiendo evaluar la integración de ambas representaciones entrenadas por separado a una red densa o equiva-

lente, o por otro lado permitiendo el ajuste de los parámetros de la red una vez integrada, utilizando los entrenamientos por separado tan solo como una inicialización de los pesos de las secciones heterogéneas y homogéneas.

En el marco de la presente iniciativa no se dispone de suficiente tiempo para abordar estas preguntas, sin embargo las baterías de experimentos realizados si demuestran que representaciones distintas al estado del arte homogéneo pueden ser exitosas en el conjunto de datos asociado al potencial fotovoltaico urbano, en particular aprovechando una mayor eficiencia computacional. Por ende, existe la posibilidad de que estas representaciones sean útiles en otros ámbitos y eventualmente puedan establecerse como una técnica más a evaluar al abordar un problema de aprendizaje automático.

Bibliografía

- [1] Swiss federal office of topography: Maps of switzerland. <https://map.geo.admin.ch/>. *Confédération Suisse*,.
- [2] Energy Technology Perspectives 2016: towards sustainable urban energy systems. *International Energy Agency*, 2016.
- [3] Ali H A Al-Waeli, K Sopian, Jabar H Yousif, Hussein A Kazem, John Boland, and Miqdam T Chaichan. Artificial neural network modeling and analysis of photovoltaic/thermal system based on the experimental study. *Energy Conversion and Management*, 186:368–379, 2019.
- [4] Eisa Almeshaie, Amin Al-Habaibeh, and Bubaker Shakmak. Rapid evaluation of micro-scale photovoltaic solar energy systems using empirical methods combined with deep learning neural networks to support systems’ manufacturers. *Journal of Cleaner Production*, 244:118788, 2020.
- [5] Khalid Anwar and Sandip Deshmukh. Assessment and mapping of solar energy potential using artificial neural network and GIS technology in the Southern part of India. *International Journal of Renewable Energy Research*, 8(2):974–985, 2018.
- [6] Dan Assouline. Machine Learning and Geographic Information Systems for large-scale mapping of renewable energy potential. Thèse N 9376, pour l’obtention du grade de Docteur ès Sciences. *École Polytechnique Fédérale de Lausanne*, 2019.
- [7] Dan Assouline, Nahid Mohajeri, and Jean-Louis Scartezzini. A Machine Learning Methodology for Estimating Roof-Top Photovoltaic Solar Energy Potential in Switzerland. *Cisbat*, (August 2016):555–560, 2015.
- [8] Dan Assouline, Nahid Mohajeri, and Jean-Louis Scartezzini. Quantifying rooftop photovoltaic solar energy potential: A machine learning approach. *Solar Energy*, 141, 01 2017.

-
- [9] Dan Assouline, Nahid Mohajeri, and Jean Louis Scartezzini. *Estimation of large-scale solar rooftop PV potential for smart grid integration: A methodological review*, volume 145. 2018.
- [10] Dan Assouline, Nahid Mohajeri, and Jean Louis Scartezzini. Large-scale rooftop solar photovoltaic technical potential estimation using Random Forests. *Applied Energy*, 217:189–211, may 2018.
- [11] A B G Bahgat, N H Helwa, G E Ahamd, and E T El Shenawy. Estimation of the maximum power and normal operating power of a photovoltaic module by neural networks. *Renewable Energy*, 29(3):443–457, 2004.
- [12] Khalil Benmouiza and Ali Cheknane. Small-scale solar radiation forecasting using ARMA and nonlinear autoregressive neural network models. *Theoretical and Applied Climatology*, 124(3-4):945–958, 2016.
- [13] J L Bosch, G López, and F J Batlles. Daily solar irradiation estimation over a mountainous area using artificial neural networks. *Renewable Energy*, 33(7):1622–1628, 2008.
- [14] Kyle Bradbury, Raghav Saboo, Timothy L. Johnson, Jordan M Malof, Arjun Devarajan, Wuming Zhang, Leslie M. Collins, and Richard G. Newell. Distributed solar photovoltaic array location and extent dataset for remote sensing object identification. *Scientific Data*, 3(1):160106, 2016.
- [15] Leo Breiman. Bagging predictors. *Machine learning*, 24(2):123–140, 1996.
- [16] Leo Breiman. Random Forests. *Machine Learning*, 45(1):5–32, 2001.
- [17] Michael Buckland and Fredric Gey. The relationship between recall and precision. *Journal of the American society for information science*, 45(1):12–19, 1994.
- [18] Pablo Campos, Lorena Troncoso, Peter D. Lund, Cristian Cuevas, Adelqui Fissore, and Rodrigo Garcia. Potential of distributed photovoltaics in urban Chile. *Solar Energy*, 135:43–49, 2016.
- [19] Christina Chatzipoulka, Raphaël Compagnon, Jérôme Kaempf, and Marialena Nikolopoulou. Sky view factor as predictor of solar availability on building façades. *Solar Energy*, 170:1026–1038, 2018.
- [20] C. Cortes and V. Vapnik. *Support-Vector Networks*, pages 273–297, 1995. cited By 144.
- [21] Christophe Paoli Cyril Voyant, Pierrick Haurant, Marc Muselli and Marie-Laure Nivet. Time series modeling and large scale global solar radiation forecasting from geostationary satellites data Cyril. *Solar Energy*, pages 1–25, 2014.
- [22] K L Du and MNS Swamy. Radial basis function networks. *Neural networks in a soft-computing framework*, pages 251–294, 2006.

-
- [23] D A Fadare. Modelling of solar energy potential in Nigeria using an artificial neural network model. *Applied Energy*, 86(9):1410–1422, 2009.
- [24] Ayoub Fentis, Lhoussine Bahatti, Mohamed Mestari, Mohamed Tabaa, Abderrahmane Jarrou, and Brahim Chouri. Short-term PV power forecasting using Support Vector Regression and local monitoring data. In *2016 International Renewable and Sustainable Energy Conference (IRSEC)*, pages 1092–1097, 2016.
- [25] Jimeno A Fonseca, Thuy-An Nguyen, Arno Schlueter, and Francois Marechal. City Energy Analyst (CEA): Integrated framework for analysis and optimization of building energy systems in neighborhoods and city districts. *Energy and Buildings*, 113:202–226, 2016.
- [26] Jerome H Friedman. *The elements of statistical learning: Data mining, inference, and prediction*. springer open, 2017.
- [27] F Ghani, M Duke, and J K Carson. Estimation of photovoltaic conversion efficiency of a building integrated photovoltaic/thermal (BIPV/T) collector array using an artificial neural network. *Solar Energy*, 86(11):3378–3387, 2012.
- [28] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-December:770–778, 2016.
- [29] T Hiyama and K Kitabayashi. Neural network based estimation of maximum power generation from PV module using environmental information. *IEEE Transactions on Energy Conversion*, 12(3):241–247, 1997.
- [30] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9:1735–80, 12 1997.
- [31] Zhaojian Huang, Thushini Mendis, and Shen Xu. Urban solar utilization potential mapping via deep learning technology: A case study of Wuhan, China. *Applied Energy*, 250:283–291, sep 2019.
- [32] Zhaojian Huang, Thushini Mendis, and Shen Xu. Urban solar utilization potential mapping via deep learning technology: A case study of Wuhan, China. *Applied Energy*, 250:283–291, 2019.
- [33] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *The Journal of Physiology*, 195:215–243, 1968.
- [34] Anwar Ibrahim and Tamer Khatib. A novel hybrid model for hourly global solar radiation prediction using random forests technique and firefly algorithm. *Energy Conversion and Management*, 138:413–425, 2017.
- [35] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *32nd International Conference on Machine Learning, ICML 2015*, 1:448–456, 2015.

-
- [36] Salvador Izquierdo, Marcos Rodrigues, and Norberto Fueyo. A method for estimating the geographical distribution of the available roof surface area for large-scale photovoltaic energy-potential evaluations. *Solar Energy*, 82(10):929–939, 2008.
- [37] Hou Jiang, Ling Yao, Ning Lu, Jun Qin, Tang Liu, Yujun Liu, and Chenghu Zhou. Geospatial assessment of rooftop solar photovoltaic potential using multi-source remote sensing data. *Energy and AI*, 10, 11 2022.
- [38] Bikash Joshi, Baluyan Hayk, Amer Al-Hinai, and Wei Woon. Rooftop Detection for Planning of Solar PV Deployment: A Case Study in Abu Dhabi. volume 8817, pages 137–149, 2014.
- [39] Jakub K Jurasz, Paweł B Dąbek, and Pietro E Campana. Can a city reach energy self-sufficiency by means of rooftop photovoltaics? Case study from Poland. *Journal of Cleaner Production*, 245:118813, 2020.
- [40] Amith Khandakar, Muhammad E. H. Chowdhury, Monzure Khoda Kazi, Kamel Benhmed, Farid Touati, Mohammed Al-Hitmi, and Antonio Jr S. P. Gonzales. Machine Learning Based Photovoltaics (PV) Power Prediction Using Different Environmental Parameters of Qatar. *Energies*, 12(14), 2019.
- [41] Tamer Khatib, Azah Mohamed, K Sopian, and M Mahmoud. Assessment of Artificial Neural Networks for Hourly Solar Radiation Prediction. *International Journal of Photoenergy*, 2012:946890, 2012.
- [42] Serkan Kiranyaz, Onur Avcı, Osama Abdeljaber, Turker Ince, Moncef Gabbouj, and Daniel J. Inman. 1d convolutional neural networks and applications: A survey. *Mechanical Systems and Signal Processing*, 151:107398, 2021.
- [43] Choongwan Koo, Taehoon Hong, Minhyun Lee, and Hyo Seon Park. Estimation of the Monthly Average Daily Solar Radiation using Geographic Information System and Advanced Case-Based Reasoning. *Environmental Science & Technology*, (9):4829–4839, 2013.
- [44] Sebastian Krapf, Nils Kemmerzell, Syed Uddin, Manuel Vázquez, Fabian Netzler, and Markus Lienkamp. Towards Scalable Economic Photovoltaic Potential Analysis Using Aerial Images and Deep Learning. *Energies*, 14:3800, 2021.
- [45] Alex Krizhevsky. ImageNet Classification with Deep Convolutional Neural Networks Alex. *Handbook of Approximation Algorithms and Metaheuristics*, pages 1–1432, 2007.
- [46] Yann Lecun and Y. Bengio. Convolutional networks for images, speech, and time-series. *The Handbook of Brain Theory and Neural Networks*, 1995.
- [47] Ziqi Li, Zidong Zhang, and Keith Davey. Estimating Geographical PV Potential Using LiDAR Data for Buildings in Downtown San Francisco. *Transactions in GIS*, 19(6):930–963, 2015.

-
- [48] Emmanuel Maggiori, Yuliya Tarabalka, Guillaume Charpiat, and Pierre Alliez. Can semantic labeling methods generalize to any city? the inria aerial image labeling benchmark. In *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, pages 3226–3229. IEEE, 2017.
- [49] Kai Mainzer, Sven Killinger, Russell McKenna, and Wolf Fichtner. Assessment of rooftop photovoltaic potentials at the urban level using publicly available geodata and image recognition techniques. *Solar Energy*, 155:561–573, 2017.
- [50] Jordan M Malof, Leslie M Collins, Kyle Bradbury, and Richard G Newell. A deep convolutional neural network and a random forest classifier for solar photovoltaic array detection in aerial imagery. In *2016 IEEE International Conference on Renewable Energy Research and Applications (ICRERA)*, pages 650–654, 2016.
- [51] A Mellit, M Benghanem, A Hadj Arab, and A Guessoum. A simplified model for generating sequences of global solar radiation data for isolated sites: Using artificial neural network and a library of Markov transition matrices approach. *Solar Energy*, 79(5):469–482, 2005.
- [52] Adel Mellit, Mohamed Benghanem, A. Hadj Arab, and Abderrezak Guessoum. Modeling of sizing the photovoltaic system parameters using artificial neural network. In *Proceedings of 2003 IEEE Conference on Control Applications, 2003. CCA 2003.*, volume 1, pages 353–357 vol.1, 2003.
- [53] Nahid Mohajeri, Dan Assouline, Agust Gudmundsson, and Jean-Louis Scartezzini. Effects of city size on the large-scale decentralised solar energy potential. *Energy Procedia*, 122:697–702, 2017.
- [54] Nahid Mohajeri, Dan Assouline, Berenice Guiboud, Andreas Bill, Agust Gudmundsson, and Jean-Louis Scartezzini. A city-scale roof shape classification using machine learning for solar energy applications. *Renewable Energy*, 121:81–93, 2018.
- [55] Nahid Mohajeri, Dan Assouline, Bérénice Guiboud, and Jean-Louis Scartezzini. DOES ROOF SHAPE MATTER? SOLAR PV INTEGRATION ON ROOFS. 2016.
- [56] M Mohandes, S Rehman, and T O Halawani. Estimation of global solar radiation using artificial neural networks. *Renewable Energy*, 14(1):179–184, 1998.
- [57] Alejandra Molina, Mark Falvey, and Roberto Rondanelli. A solar radiation database for Chile. *Scientific Reports*, 7(1):1–11, 2017.
- [58] A Moreno, M A Gilabert, and B Martínez. Mapping daily global solar irradiation over Spain: A comparative study of selected approaches. *Solar Energy*, 85(9):2072–2084, 2011.
- [59] Guy Newsham and Benjamin Birt. Building-level occupancy data to improve arima-based electricity use forecasts. *BuildSys’10 - Proceedings of the 2nd ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings*, 01 2010.

-
- [60] S Premrudeepreechacharn and N Patanapirom. Solar-array modelling and maximum power point tracking using neural networks. In *2003 IEEE Bologna Power Tech Conference Proceedings*, volume 2, pages 5 pp. Vol.2–, 2003.
- [61] Prajit Ramachandran, Niki Parmar, Ashish Vaswani, Irwan Bello, Anselm Levskaya, and Jonathon Shlens. Stand-alone self-attention in vision models. *CoRR*, abs/1906.05909, 2019.
- [62] K S Reddy and Manish Ranjan. Solar resource estimation using artificial neural networks and comparison with other correlation models. *Energy Conversion and Management*, 44(15):2519–2530, 2003.
- [63] Haoshan Ren, Chengliang Xu, Zhenjun Ma, and Yongjun Sun. A novel 3d-geographic information system and deep learning integrated approach for high-accuracy building rooftop solar energy potential characterization of high-density cities. *Applied Energy*, 306, 1 2022.
- [64] Raúl Rojas. *The Backpropagation Algorithm*, pages 149–182. Springer Berlin Heidelberg, Berlin, Heidelberg, 1996.
- [65] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.
- [66] Meita Rumbayan, Asifujiang Abudureyimu, and Ken Nagasaka. Mapping of solar energy potential in Indonesia using artificial neural network and geographical information system. *Renewable and Sustainable Energy Reviews*, 16(3):1437–1449, 2012.
- [67] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv 1409.1556*, 09 2014.
- [68] Yeliz Simsek, Álvaro Lorca, Tania Urmee, Parisa A. Bahri, and Rodrigo Escobar. Review and assessment of energy policy developments in Chile. *Energy Policy*, 127(November 2018):87–101, 2019.
- [69] Adnan Sözen, Erol Arcakliolu, and Mehmet Özalp. Estimation of solar potential in Turkey by artificial neural networks using meteorological and geographical data. *Energy Conversion and Management*, 45(18):3033–3052, 2004.
- [70] Adnan Sözen, Erol Arcaklolu, Mehmet Özalp, and Naci Çalar. Forecasting based on neural network approach of solar potential in Turkey. *Renewable Energy*, 30(7):1075–1090, 2005.
- [71] Huaiwei Sun, Dongwei Gui, Baowei Yan, Yi Liu, Weihong Liao, Yan Zhu, Chengwei Lu, and Na Zhao. Assessing the potential of random forest method for estimating solar

-
- radiation using air pollution index. *Energy Conversion and Management*, 119:121–129, 2016.
- [72] Tao Sun, Ming Shan, Xing Rong, and Xudong Yang. Estimating the spatial distribution of solar photovoltaic power generation potential on different types of rural rooftops using a deep learning network applied to satellite images. *Applied Energy*, 315, 6 2022.
- [73] Lisa Torrey and Jude Shavlik. Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques*, pages 242–264. IGI global, 2010.
- [74] US EIA. International Energy Outlook 2013 - DOE/EIA-0484(2013). *Outlook 2013*, page 312, 2013.
- [75] Alvaro Valderrama, Carlos Valle, Hector Allende, Marcelo Ibarra, and Camila Vásquez. Machine learning applications for urban photovoltaic potential estimation: A survey. *Neurocomputing*, 526:80–95, 2023.
- [76] Alvaro Valderrama, Carlos Valle, Marcelo Ibarra, and Héctor Allende. A heterogeneous 1d convolutional architecture for urban photovoltaic estimation. In *Intelligent Computing Theories and Application - 17th International Conference, ICIC 2021, Shenzhen, China, August 12-15, 2021, Proceedings, Part I*, volume 12836 of *Lecture Notes in Computer Science*, pages 435–449. Springer, 2021.
- [77] Aristotelis Vartholomaios. A machine learning approach to modelling solar irradiation of urban and terrain 3D models. *Computers, Environment and Urban Systems*, 78:101387, 2019.
- [78] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *CoRR*, abs/1706.03762, 2017.
- [79] Alina Walch, Roberto Castello, Nahid Mohajeri, Fabian Guignard, Mikhail Kanevski, and Jean-Louis Scartezzini. Spatio-temporal modelling and uncertainty estimation of hourly global solar irradiance using Extreme Learning Machines. *Energy Procedia*, 158:6378–6383, 2019.
- [80] Alina Walch, Roberto Castello, Nahid Mohajeri, and Jean-Louis Scartezzini. Big data mining for the estimation of hourly rooftop photovoltaic potential and its uncertainty. *Applied Energy*, 262:114404, 2020.
- [81] L. K. Wiginton, H. T. Nguyen, and J. M. Pearce. Quantifying rooftop solar photovoltaic potential for regional renewable energy policy. *Computers, Environment and Urban Systems*, 34(4):345–357, 2010.
- [82] WenBo Xiao, Gina Nazario, HuaMing Wu, HuaMing Zhang, and Feng Cheng. A neural network based computational model to predict the output power of different types of photovoltaic cells. *PLOS ONE*, 12(9):1–8, 2017.

-
- [83] Amit Kumar Yadav and S S Chandel. Solar energy potential assessment of western Himalayan Indian state of Himachal Pradesh using J48 algorithm of WEKA in ANN based prediction model. *Renewable Energy*, 75:675–693, 2015.
- [84] Xin-She Yang. Firefly Algorithms for Multimodal Optimization. In Osamu Watanabe and Thomas Zeugmann, editors, *Stochastic Algorithms: Foundations and Applications*, pages 169–178, Berlin, Heidelberg, 2009. Springer Berlin Heidelberg.
- [85] Xin-She Yang. *Nature-Inspired Metaheuristic Algorithms*. 2010.
- [86] Chen Zhang, Zhixin Li, Haihua Jiang, Yongqiang Luo, and Shen Xu. Deep learning method for evaluating photovoltaic potential of urban land-use: A case study of Wuhan, China. *Applied Energy*, 283:116329, 2021.
- [87] Teng Zhong, Kai Zhang, Min Chen, Yijie Wang, Rui Zhu, Zhixin Zhang, Zixuan Zhou, Zhen Qian, Guonian Lv, and Jinyue Yan. Assessment of solar photovoltaic potentials on urban noise barriers using street-view imagery. *Renewable Energy*, 168:181–194, 5 2021.
- [88] Teng Zhong, Zhixin Zhang, Min Chen, Kai Zhang, Zixuan Zhou, Rui Zhu, Yijie Wang, Guonian Lü, and Jinyue Yan. A city-scale estimation of rooftop solar photovoltaic potential based on deep learning. *Applied Energy*, 298, 9 2021.
- [89] Guoqing Zhou, Weirong Chen, John A Kelmelis, and Deyan Zhang. A comprehensive study on urban true orthorectification. *IEEE Transactions on Geoscience and Remote sensing*, 43(9):2138–2147, 2005.
- [90] Hangxia Zhou, Yujin Zhang, Lingfan Yang, Qian Liu, Ke Yan, and Yang Du. Short-Term Photovoltaic Power Forecasting Based on Long Short Term Memory Neural Network and Attention Mechanism. *IEEE Access*, 7:78063–78074, 2019.
- [91] Ümit Abulut, Ali Etem Gürel, Alper Ergün, and Ihan Ceylan. Performance assessment of a V-trough photovoltaic system and prediction of power output with different machine learning algorithms. *Journal of Cleaner Production*, 268:122269, 2020.