**El trabajo doméstico no remunerado: una oportunidad para la focalización de políticas**

**Carlos A. Ayala**

**Mateo Contreras**

**Federico Meneses**

**Abstract**

El trabajo no remunerado, considerando especialmente aquellas actividades relacionadas con el cuidado del hogar, ha tomado día a día más importancia, y con ello la creación de políticas para equilibrar las disparidades de tiempo de trabajo no remunerado entre mujeres y hombres. El objetivo de este documento es proponer un modelo de predicción que permita focalizar políticas, servicios, protecciones sociales e infraestructura que promueva la diminución del tiempo ejercido por las mujeres en trabajo no remunerado .Para este documento nos enfocaremos en la predicción del tiempo que se dedica un hombre o una mujer en un hogar en base a datos obtenidos del Dane y por medio de un modelo XGBoost (Extreme Gradient Boosting) se entrena y predice un modelo que estima este tiempo para los hombres y mujeres de ese hogar y luego se comparan las brechas de los mismos. Una brecha muy alta implicaría que es un hogar que necesita las políticas, sin embargo, se considera que esto ocurrirá para algunos y no para todos. Los resultados predicen que todos los hogares necesitan de política pública contrario a lo considerado previamente.

**Introducción**

El trabajo doméstico no remunerado es aquella actividad que no devenga un salario y compete al cuidado de niños o adultos mayores, junto a las actividades de cuidado del hogar. Estas actividades no contables dentro de una economía, y presentan un grave problema en asignación del trabajo, siendo las mujeres quienes asumen en promedio la mayor. Según estimados de ONUWOMEN (s.f.), este tipo de trabajo representa hasta un 39% del PIB, en donde las mujeres dedican poco más del doble del tiempo que dedican los hombres para las mismas labores. Por lo anterior, se hacer relevante afrontar la siguiente pregunta, ¿es posible, por medio de características observables de las personas, predecir la cantidad de horas que le dedicará al trabajo no remunerado? Con ello, es posible ahondar en la necesidad de promover acciones que puedan equilibrar esta dedicación laboral entre hombres y mujeres de forma eficiente. En concordancia con lo anterior, se explora la efectividad de un modelo de predicción para poder mejorar la contabilidad nacional sobre tiempo de trabajo no remunerado, y que esta sirva de insumo a la hora de la asignación de políticas públicas enfocadas en disminuir las brechas existentes. Consideramos que un modelo de predicción para esta situación podría reducir costos en recolección de información además colocar eficientemente lo hogares que requieran estas políticas.

Nuestro proyecto se diferencia de otros varios dado que la literatura que se encuentra alrededor del tema es literatura inclinada únicamente a describir el trabajo no remunerado o tratar de explicar sus efectos sobre el PIB. Se han encontrado diferentes artículos abordan esta situación y buscan explicar la razón de esta problemática. Por un lado, Zachorowska-Mazurkiewicz (2016) que argumenta como economistas neoclásicos como Pigou o Marshall concordaban en no contabilizar el valor agregado que genera este trabajo doméstico no remunerado en las cuentas nacionales. Por otro lado, se explica que la limpieza era una tarea moralmente inherente a las mujeres (Ehrenreich y English,1976, como se cita en Hazarika y Das, 2021). Estas declaraciones incentivaron más el enfoque de las tareas domésticas a las mujeres hasta que se tomó como un trabajo de tiempo completo, esto las mantuvo fuera del mercado laboral. Con la creciente participación de las mujeres en el mercado laboral, se considera importante establecer qué población tiene mayor dedicación al trabajo doméstico no remunerado y promover políticas públicas que puedan ayudar a equilibrar la carga laboral y mejorar el bienestar de la población.

Se considera que los resultados obtenidos serán acordes a la literatura, es decir, se logrará predecir que, para la mayoría de las mujeres, el tiempo de dedicado al trabajo no remunerado alrededor del 100% que la contraparte masculina le dedica a estos trabajos. Sin embargo, es necesario destacar que el objetivo está ligado es a obtener predicciones acertadas sobre el tiempo trabajado por hombres y mujeres dadas las características de un hogar y comparar si las brechas requieren una aplicación eficaz de políticas públicas relacionadas con subsidios, servicios, educación o demás programas que se consideren para la reducción de estas brechas, en caso de que las brechas de tiempo que se presenten sean que las mujeres trabajen el doble de tiempo o más que los hombres.

**Data**

Colombia, a través del Departamento Administrativo Nacional de Estadísticas DANE, ha realizado seguimiento cuidadoso a la distribución del uso del tiempo en el territorio nacional por medio de la Encuesta Nacional de Uso del Tiempo ENUT. La encuesta reúne información a nivel de hogares como condiciones de vivienda y características como su composición y condiciones socioeconómicas. Adicionalmente, la encuesta contiene información desagregada por individuo como condición de salud, tiempo dedicado a crianza y cuidado de niños menores de 5 años, educación, sexo, condiciones laborales, tiempo dedicado a actividades diferentes a crianza de niños menores a 5 años e información sobre cambios en las actividades realizadas en el contexto de la emergencia sanitaria por COVID-19.

La base de datos constaba de 9 planillas de Excel con diferentes variables, pero todas con un mismo identificador por el cual es posible unir los data frames. Estas variables son “DIRECTORIO”, “SECUENCIA\_P” y “ORDEN”, sin embargo, para nuestro análisis cambiamos sus nombres respectivamente a vivienda hogar e ind (individuo). La selección de las variables a utilizar estuvo dada por características del hogar, tales como si posee servicios, salud, educación, electrodomésticos, si alguien del hogar sufre crónicamente de alguna enfermedad, nivel educativo máximo alcanzado por el hogar o cuantos miembros tienen el hogar; se tiene de la misma forma características de los individuos, como sexo, raza, estado civil o si vivió o no en el campo.

Ahora, por otra parte, el tiempo de trabajo no remunerado en el hogar, se obtuvo mediante la sumatoria de las variables la página C\_8 del ENUT. Entre estas se encontraba el tiempo que se gasta en diferentes labores de cocina, reparaciones, administración de la casa, limpieza, y mantenimiento de vestuario. Esta nueva variable generada es un agregado total en minutos de estas diferentes actividades que se efectúan en el hogar y que son no remuneradas. Finalmente, para obtener la base final, se une la nueva variable tiempo\_labor\_no\_rem a la base de datos con las variables de control. Una vez se une el vector con la base de datos, se utiliza la categórica de sexo para obtener dos bases de datos, una con las variables y el tiempo de los hombres y otra con las variables y tiempo de las mujeres. Es necesario resaltar que al efectuar este proceso también se eliminaron los “missing values” de la base de datos, por lo que obtenemos dos bases cada una de aproximadamente 28000 observaciones, por lo que más adelante no se imputan datos.

Ahora, en *tabla.1* se observan las estadísticas descriptivas para las subbase de datos generada previamente de mujeres. Es posible evaluar esta tabla por secciones, ignorando los identificadores, la primera sección es posible observar si la mujer posee servicios o no, en su mayoría se puede confirmar que las mujeres, en promedio, más del 70% posee todos los servicios, siendo más bajo el de internet. La siguiente sección importante sería si posee un electrodoméstico en particular o no, se destaca el hecho que en promedio menos del 20% de las mujeres poseen horno, lava platos y secadora, sin embargo, para los demás electrodomésticos, más del 50% de las mujeres los posee. En promedio las mujeres de esta encuesta tienen 44 años, no es posible observar a que grupo étnico pertenecen dado que esto categorizará después en el XGBoost, así como Estado civil. Más del 95% tienen pareja y menos del 5% viven con su padre madre. La variable de interés, tiempo de labores no remuneradas, indicad que aproximadamente en promedio las mujeres gastan alrededor de 250 minutos, lo equivalente a 4 horas al día.

Por otro lado en *tabla.2* , que muestra las estadísticas descriptivas de la sub base hombres, resulta tener promedios muy similares en casi todas las características a las mujeres, cuyas únicas diferencias relevantes están en el uso de bicicletas y automóviles, que tienen unos puntos porcentuales más altos en parejas que las mujeres y que en promedio se dedican en promedio alrededor de 60 minutos al día para los trabajos del hogar no remunerados, es decir, 3 horas menos que las mujeres.Estas tablas nos dan entonces indicios que, pese a tener características similares tanto individuales como en el hogar, lo hombres dedican en promedio 3 horas menos que las mujeres en labores del hogar, es decir, 3 veces menos.

**Modelo**

Para este documento se decidió utilizar un XGBoost. Este modelo en particular funciona apartir de ser un resampleo aleatorio de árboles con penalizaciones frente al número de iteraciones que se resamplee. Este modelo no tiene interpretabilidad teórica para, por ejemplo, entender qué factores o variables dentro del modelo resultan ser más importantes, o cómo las interacciones entre estas pueden afectar la predicción. Sin embargo, para nuestros objetivos, esto no es necesario, pues únicamente queremos evaluar un método de predicción de las brechas entre géneros para aplicar más focalizada mente una política. Lo que sí va a ser necesario para el XGBoost es tener en cuenta la selección de hiperparámetros y en especial el número de rondas las cuales vamos a iterar el número de árboles a generar y tratar. Para decidir este hiperparámetro, es posible utilizar el método de validación cruzada, el cual nos permite observar para diferentes valores de rondas los resultados del modelo y más importante aún, el error cuadrático medio de cada modelo generado en la iteración.

Entonces, se decide utilizar un XGBoost con validación cruzada la cual me decide uno de los hiperparámetros, sin embargo, cómo se seleccionan todos los demás. Otros hiperparámetros como tasas de aprendizaje, gamma (controlador del ajuste) y máxima profundidad del árbol, van a ser las predeterminadas de R, esto debido a que, para estos parámetros, las diferencias son marginales y así mismo facilita la replicabilidad del modelo.

El criterio bajo el cual se va a elegir entonces el mejor XGBoost para mis predicciones, va a ser por medio del previamente mencionado error cuadrático medio, este representa la varianza que se obtienen de los estimadores obtenidos en el modelo en los datos de entrenamiento para su posterior uso en una sección de datos de testeo, Un error cuadrático medio pequeño entonces implica estimadores con muy poca varianza y así mismo predicciones muchos más acertadas fuera de muestra.

Una vez explicado el modelo, es necesario explicar qué se va a predecir con él. Se busca entonces predecir los estimadores para los modelos con los tiempos de labor no remunerada de hombre y de mujer por aparte, es decir, se generan dos modelos de predicción del tiempo para comprobar específicamente las predicciones para cada uno por aparte.

**Resultados:**

Para observar con exactitud las predicciones es posible abrirlas en la documentación de GitHub, se llaman “thatm\_xg.csv” y “thath\_xg.csv". Para efectos prácticos de los resultados, analizaremos lo brindado por los dos modelos escogidos en la *tabla.3*. Los dos modelos elegidos fueron aquellos que menor error cuadrático medio para subbase de datos, para las mujeres el error cuadrático medio fue de 16512.18, mientras que para los hombres fue de 7876.5. no se coloca el número de rondas obtenidos pero su reproducción será la misma si se aplica la misma seed que se usó y que está con los códigos antes de cada código de XGBoost.

Al remitirse a la *tabla.3* es posible observar que la media de los datos predichos, los yhat h y m, tienen el mismo promedio y se acercan bastante a los promedios anteriormente obtenidos en las estadísticas descriptivas, esto parece entonces indicar que hay cercanía entre las predicciones y los valores reales, sin embargo, la desviación estándar varía y es mucha más pequeña a la original. También se obtuvieron algunos valores negativos indicados en los percentiles. Los percentiles para ambos géneros parecen indicar también que las predicciones resultaron estar bastante cercanas a la media, esto para el caso de hombres como de mujeres, lo que indica que los estimadores de las predicciones pueden estar subestimando los valores reales.

Por parte de la brecha, la brecha resulta ser un resultado esperado, pues como se observó en la media de las estadísticas descriptivas para el tiempo dedicado a las labores del hogar de cada género, resultaron tener una brecha en promedio del triple del tiempo entre hombres y mujeres. Esto se refleja en que más del 95 por ciento de los casos se presenten brechas del doble o más entre los tiempos dedicados al hogar de cada sexo. Esto indica una desigualdad enorme entre los tiempos dedicados al hogar entre hombres y mujeres de los hogares colombianos de forma casi que generalizada para todos los hogares.

**Conclusiones:**

El trabajo no remunerado, considerado como labores domésticas y de cuidado, ha demostrado ser un problema global y generalizado tal cómo lo especifican organismos como las Naciones Unidas. Dicha problemática fue abordada en este trabajo por medio de un modelo XGBoost de predicción, con el cual se encontró que en la mayoría de los hogares, exisitía una alta probabilidad (96%) de encontrar una brecha de tiempo entre hombres y mujeres del más del doble. Los resultados se encuentran cosistentes a lo largo de los diferentes modelos planteados.

Por lo anterior, se concluye que la reducción de las brechas de tiempo de trabajo no remunerado en el caso colombiano requiere de políticas pensadas en un ámbito general. Por tanto, considerar políticas focalizadas por medio de programas de transferencias y asistencia y cobertura de salud, no significan un efecto redistribuidor sobre la asignación de horas de trabajo no remunerado entre hombres y mujeres. Los resultados del modelo sugieren que estos programas se asocian con la existencia de brechas en trabajo no remunerado de más del doble de tiempo para el 95% de los hogares observados.

**Data viability:**

[**https://github.com/CarlosAyala123/Proyecto-final-Prediccion-uso-del-tiempo**](https://github.com/CarlosAyala123/Proyecto-final-Prediccion-uso-del-tiempo)

1. **Referencias & Bibliografía:**

*Referencias:*

ONUWOMEN. (s.f.). Redistribuir el trabajo no remunerado. Recuperado de: <https://www.unwomen.org/es/news/in-focus/csw61/redistribute-unpaid-work>

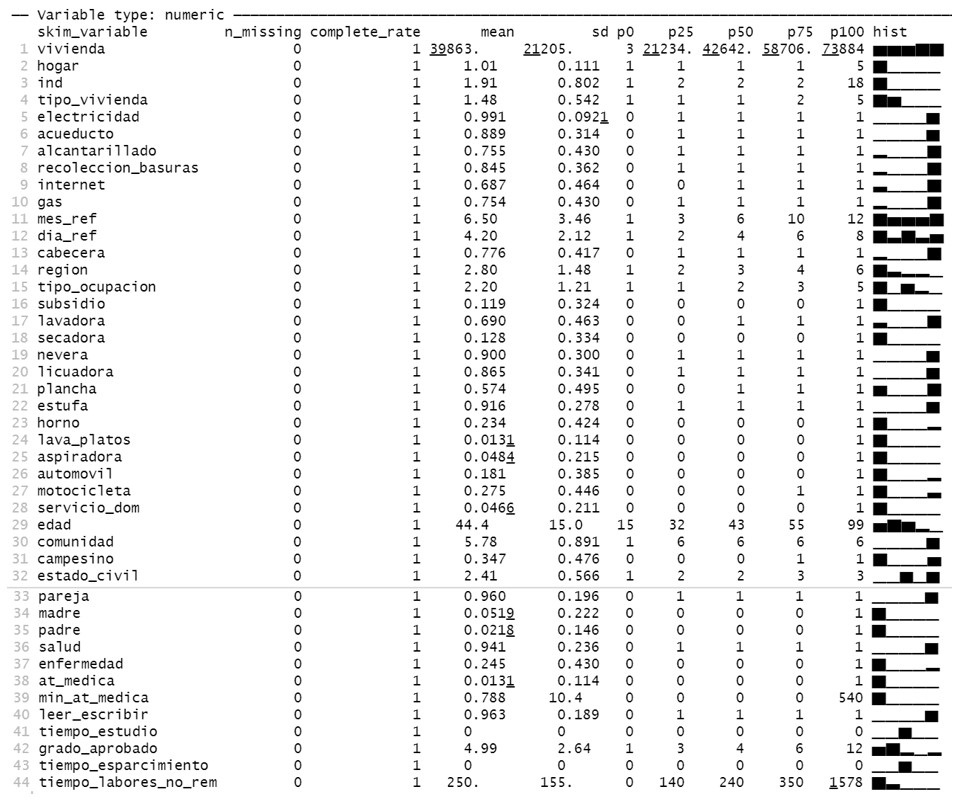
Zachorowska-Mazurkiewicz, A. (2016). GENDER, UNPAID LABOUR AND ECONOMICS. Recuperado de: <https://dspace.uni.lodz.pl/bitstream/handle/11089/21763/0_2_%5b121%5d-132_Zachorowska-Mazurkiewicz.pdf?sequence=1&isAllowed=y>

Obja Borah Hazarika & Sarmistha Das (2021) Paid and unpaid work during the Covid-19 pandemic: a study of the gendered division of domestic responsibilities during lockdown, Journal of Gender Studies, 30:4, 429-439, DOI: [10.1080/09589236.2020.1863202](https://doi.org/10.1080/09589236.2020.1863202)

*Bibliografía:*

Ramos, L. (2021). Labores de cuidado y trabajo doméstico no remunerado. Recuperado de: <https://onuhabitat.org.mx/index.php/labores-de-cuidado-y-trabajo-domestico-no-remunerado>

**Anexos:**

* *Tabla.1*****
* *Tabla.2***Tabla

  Descripción generada automáticamente**
* *Tabla.3***Calendario

  Descripción generada automáticamente**