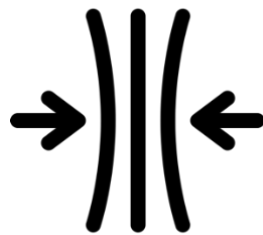


ESTUDIO DE LA RESILIENCIA

Investigación sobre el poder predictivo de
características personales y de contexto en la
capacidad de resiliencia de una persona

TRABAJO DE FIN DE MÁSTER
MÁSTER DE DATA SCIENCE
ESCUELA KSCHOOL



ÍNDICE

1. Introducción	pág. 3
2. Descripción del dataset	pág. 5
a. Caracterización sociodemográfica	pág. 5
b. Caracterización de relaciones interpersonales	pág. 6
c. Caracterización contextual COVID-19	pág. 6
d. Caracterización psicológica	pág. 6
3. Metodología	pág. 9
a. Exploración, análisis y selección de la información	pág. 9
b. Modelado	pág. 10
c. Interpretación	pág. 11
d. Desarrollo de webapp	pág. 12
4. Resumen de resultados	pág. 13
5. Conclusiones	pág. 15
6. Manual de usuario para la aplicación interactiva	pág. 16

1. Introducción

Recientemente, desde que tuvo lugar en 2020 la pandemia global del COVID-19 y el confinamiento y la alteración de las costumbres sociales asociados a la misma, ha habido un interés creciente en la sociedad por un concepto ya conocido y utilizado en el campo de la psicología desde los años 70 del siglo pasado: la resiliencia.

A pesar de llevar muchos años bajo estudio, se trata de un término con cierta controversia, ya que no existe un consenso entre los expertos sobre su definición última. El concepto de resiliencia está sujeto a diferentes interpretaciones y por tanto es difícil medirlo de manera precisa mediante una escala psicológica, ya que un constructo que para algunos investigadores resulte completo, no será sin embargo suficiente para otros.

Para este estudio se utiliza la definición de uno de los investigadores más relevantes en los últimos años en el campo, el Dr. George Bonanno, que define resiliencia como un funcionamiento saludable y estable de la persona ante un evento potencialmente traumático.

Quizás el aspecto que más dificultad entraña para crear una métrica fiable para la resiliencia consista en su dependencia del contexto: el momento y ante qué evento que se mide. Las investigaciones recientes se han centrado en intentar crear esa métrica mediante cuestionarios auto-informados que tratan de concretar las capacidades de la persona a la hora de afrontar y adaptarse a dificultades.

Otro enfoque diferente consistiría en medir la resiliencia de manera empírica a través del impacto que tiene el contexto que está viviendo la persona en su salud mental: si aumenta sintomatología relacionada con estrés, ansiedad o depresión, y se da además interferencia con su capacidad de funcionar diariamente, podríamos resolver que la persona no está siendo resiliente.

Esta línea de investigación es la que aborda un equipo liderado por la Dra. Carmen Valiente y el Dr. Carmelo Vázquez, del departamento de Psicología Clínica de la Universidad Complutense de Madrid, cuando se unieron a un Consorcio internacional formado por diversas universidades europeas para estudiar el efecto en la salud mental de la pandemia COVID-19. Diseñaron un estudio longitudinal para medir los cambios psicológicos que se producían en la población general adulta española a lo largo de la duración de la pandemia. Se hizo mediante una encuesta a través de internet y lanzada en dos momentos diferentes: T1, tuvo lugar durante el pico de la pandemia (7-14 de abril 2020) en medio del confinamiento y T2, justo antes de que comenzaran las medidas de relajación de dicho confinamiento (7-11 de mayo 2020), cuando las muertes diarias ya estaban disminuyendo notablemente. De esta manera, es posible perfilar a la muestra en T1 y conocer su estado en T2, extrayendo un valor de resiliencia de cuatro posibles (“Resiliente”, “Recuperado”, “Distrés sostenido”, “Distrés retrasado”) para cada observación (participante).

Su estudio investigó el papel de variables sociodemográficas, de salud, psicológicas y de relaciones interpersonales en las anteriores categorías de resiliencia, encontrando mediante un modelo de regresión logística multinomial que una mayor probabilidad de resiliencia estaba asociada con ser hombre, tener más edad, no tener antecedentes de

salud mental, y tener altos niveles de identificación con la humanidad y de bienestar psicológico. Además, bajos niveles en otras variables (ansiedad y amenaza económica debida a la pandemia, consumo de sustancias durante el confinamiento, intolerancia a la incertidumbre, ansiedad ante la muerte, soledad, y desconfianza) fueron también predictores de una respuesta de resiliencia psicológica al COVID-19.

El objetivo general de este TFM es ampliar el estudio descrito anteriormente, utilizando otras técnicas de aprendizaje automático. En concreto, los objetivos son:

- Analizar los factores con mayor poder predictivo y validar/ampliar los hallazgos del estudio anterior.
- Crear un modelo predictivo de resiliencia para la población general dependiente sólo de características personales, es decir, independiente del contexto.
- Desarrollar una aplicación interactiva basada en una simplificación del modelo anterior que pueda servir como herramienta para testar el nivel de resiliencia de una persona en un futuro cercano ante un contexto difícil.

En los últimos años ha habido un incremento de las investigaciones sobre el uso de métodos de aprendizaje automático en el campo de la psicología como complemento a técnicas de diagnóstico y selección de posibles tratamientos terapéuticos. Este trabajo entra dentro de este campo de investigación y pretende aportar nueva evidencia sobre los aspectos de la metodología de diagnóstico y tratamiento que son susceptibles de incrementar su eficacia con el uso de técnicas de aprendizaje automático.

2. Descripción del dataset

El dataset utilizado es la encuesta mencionada en el estudio del equipo de la UCM. La encuesta, repetida en dos momentos diferentes (segunda semana de abril de 2020 y segunda semana de mayo de 2020), estaba compuesta por variables pertenecientes a diversos ámbitos que caracterizan a una persona:

- Caracterización sociodemográfica y de condiciones de vida
- Caracterización contextual COVID-19
- Caracterización psicológica
- Caracterización de relaciones interpersonales

Aunque esta encuesta es recibida procesada, es decir, con las métricas totales de cada escala construidas, también incluye los diferentes ítems individuales que las componen. Dado el enorme número de variables que contiene el dataset, vamos a describir aquí tan sólo las variables que quedan tras la selección realizada durante el análisis exploratorio que se puede seguir en el notebook llamado 01_dataselection.ipynb y que se vuelcan sobre el fichero covid_dataset_csv. En este notebook se puede encontrar un análisis de cada variable. Como referencia, se puede encontrar también en el repositorio el documento covid19_survey.pdf que contiene las preguntas de la encuesta.

Caracterización sociodemográfica

'Edad'

- Edad

'Sexo'

- Sexo

'Zona_vivienda'

- Tipo de área en la que se ubica la vivienda del participante

'Pertenencia_vecindario'

- Sentimiento de pertenencia al vecindario

'Comodidad_vecindario_llaves'

- Integración con el vecindario (dejaría llaves)

'Comodidad_vecindario_compra'

- Integración con el vecindario (le harían la compra)

'Titulación'

- Nivel de estudios

'Situación_laboral'

- Situación laboral

'Adultos_vivienda'

- Número de adultos en la vivienda

'Menores_vivienda'

- Número de menores en la vivienda

'Descripción_lugar_vivienda'

- Nivel de propiedad de la vivienda

'Ingresos'

- Nivel de ingresos

'pérdida_ingresos'

- Si ha sufrido pérdida de ingresos por la pandemia

'SUBSTANCE_SCORE'

- Incremento de consumo de sustancias

'Religión'

- Creencia religiosa

'Ideología_politica'

- Ideología política (derechas – centro – izquierda)

'PatriotismoTOTAL'

- Grado de sentimiento patriótico

'NacionalismoTOTAL'

- Grado de sentimiento nacionalista

'Authoritharism_VSAS'

- Very short authoritarianism scale (VSAS), mide el grado de actitud autoritaria frente a liberal

'RELIGIOUSITY'

- Grado de religiosidad

Caracterización de relaciones interpersonales

'SDO_TOTAL'

- SOCIAL DOMINANCE ORIENTATION SCALE (SDO), mide el grado de sentimiento de dominancia social

'ATTITUDE_MIGRANTS'

- Grado de aceptación de la inmigración

'HUMANITY_TOTAL'

- Grado de identificación con valores humanos

'Teorías_de_la_conspiración'

- Grado de sentimiento de conspiración social

'CONFIANZA_EN_OTROS'

- Grado de confianza en otras personas

'LONELI'

- Three-Item Loneliness Scale (TILS), mide el sentimiento de soledad y aislamiento

Caracterización contextual COVID-19

'AISLAMIENTO'

- Si ha estado aislado por contagio de COVID-19

'preocupación_economía_covid'

- Grado de preocupación por la economía tras la pandemia

'ANSIEDAD_POR_COVID'

- Si ha tenido síntomas de ansiedad debido a la pandemia

'COVID_KNOWLEDGE'

- Grado de conocimiento general sobre el COVID-19

'COVID_TRANSMISSION'

- Grado de conocimiento sobre mecanismos de contagio del COVID-19

'COVID_ENGAGEMENT'

- Grado de conocimiento sobre modo de actuar si contagio de COVID-19

'VACUNA_USTED'

- Si se vacunaría en caso de desarrollo de vacuna

'JTC_COVID'

- Respuesta ante situaciones COVID-19

Caracterización psicológica

'Saludmental_tratamiento_pasado'

- Si ha tenido tratamiento de salud mental en el pasado

'Saludmental_tratamiento_presente'

- Si tiene tratamiento de salud mental en el presente

'PHQ_15_TOTAL'

- Patient Health Questionnaire-15 (PHQ-15), mide la gravedad de síntomas somáticos durante la última semana

'PHQ_9_TOTAL'

- Patient Health Questionnaire-9 (PHQ-9), mide la gravedad de síntomas depresivos durante las dos últimas semanas

'DEP_DIAGN'

- Variable derivada que diagnostica depresión a partir de respuestas en PHQ_9_TOTAL (si es ≥ 5) y si la respuesta a los dos primeros ítems es afirmativa)

'GAD_TOTAL'

- Generalized Anxiety Disorder Scale (GAD-7), mide la gravedad de síntomas de ansiedad en las últimas dos semanas

'PADS_TOTAL'

- Short-form Persecution and Deservedness Scale (SF-PaDS), mide la gravedad de pensamientos de ideación paranoide

'PTSD_REEXP'

- Subescala del Posttraumatic Stress Disorder Checklist (PCL-5), mide la reexperimentación del trauma

'PTSD_AVOID'

- Subescala del Posttraumatic Stress Disorder Checklist (PCL-5), mide la evitación del trauma

'PTSD_THREAT'

- Subescala del Posttraumatic Stress Disorder Checklist (PCL-5), mide la amenaza del trauma

'PTSD_SEVERITY'

- Subescala del Posttraumatic Stress Disorder Checklist (PCL-5), mide la gravedad del trauma

'PTSD_IMPAIR'

- Subescala del Posttraumatic Stress Disorder Checklist (PCL-5), mide la interferencia en la vida del trauma

'PTSD_DIAGNOSIS'

- Variable derivada de las subescalas del PCL-5, mide el diagnóstico del trauma

'PTSD_TOTAL'

- Posttraumatic Stress Disorder Checklist (PCL-5), mide el grado de afectación de estrés posttraumático

'SUBST_STRESS'

- Variable derivada de las subescalas del PCL-5 para diagnosticar estrés sostenido

'CRT_TOTAL'

- Cognitive Reflection Test (CRT), mide la tendencia a anular una respuesta visceral incorrecta y a reflexionar más para encontrar una respuesta correcta

'LMLOS_chance'

- Subescala del Levenson Multidimensional Locus of Control Scale (LMLOS), mide la sensación de influencia del azar en los acontecimientos la vida

'LMLOS_powerful'

Estudio de la Resiliencia

- Subescala del Levenson Multidimensional Locus of Control Scale (LMLOS), mide la sensación de influencia de otros en los acontecimientos la vida

'LMLOS_internal'

- Subescala del Levenson Multidimensional Locus of Control Scale (LMLOS), mide la sensación de influencia propia en los acontecimientos la vida

'OFS_total'

- Openness to the Future Scale (OFS), mide la actitud positiva de la persona hacia el futuro

'BRS_total'

- Brief Resilience Scale (BRS), mide la auto capacidad percibida de recuperarse de las dificultades

'DAI_TOTAL'

- Death Anxiety Inventory (DAI), mide la ansiedad hacia la muerte

'IUS_total'

- Intolerance to Uncertainty Scale (IUS), mide la intolerancia hacia la incertidumbre

'PEMBERTON_TOTAL'

- Pemberton Happiness Index (PHI), mide el bienestar de la persona

'PI_total'

- Primals Inventory (PI), mide las creencias y valores primarios de la persona

'PGT_SF_total'

- Posttraumatic Growth Inventory-Short Form (PGTI-SF), mide el grado en que los supervivientes de sucesos traumáticos perciben los beneficios personales que se derivan de sus intentos de hacer frente al trauma y sus secuelas

'JTC_GENERAL'

- Respuesta ante situaciones generales

'RESILIENT_W1234'

- Número de 4 cifras, una por cada ola de encuestación. Cada una de ellas puede tomar 3 valores: 0 (no resiliente), 1 (resiliente) y 9 (no disponible).
- La asignación de valor para cada participante en cada ola depende de si las observaciones para PHQ-9, GAD-7 y PCL-5 alcanzan unos valores umbral.
- Es por tanto una asignación basada en síntomas empíricos relacionados con la depresión, la ansiedad y el estrés.
- Utilizamos sólo las dos primeras cifras y recodificaremos de la siguiente manera:
 - o Resilient: 00
 - o Sustained distress: 11
 - o Delayed distress: 01
 - o Recovered: 10

3. Metodología

Para cumplir con el objetivo principal del estudio detallado en el punto 1, la estrategia general seguida se basa en entrenar un modelo de aprendizaje automático clasificador que utilice como target el estado de resiliencia empírico (ver última variable del punto anterior) en la segunda ola, y como features las variables caracterizadoras del participante tomadas en la primera ola.

De esta manera, estaremos prediciendo el estado de resiliencia de una persona a un mes vista de conocer sus características sociodemográficas, interpersonales y psicológicas, en un contexto que supone un reto o dificultad potencialmente traumáticos para ella.

El entorno de trabajo escogido está basado en notebooks con Jupyter sobre el lenguaje Python.

Exploración, análisis y selección de la información

Los pasos para analizar y seleccionar la información con la que alimentar al modelo se incluyen en el notebook llamado 01_dataselection. En concreto, se realizaron las tareas siguientes:

- Comprobación y estudio del dataset
- Selección de features de las primeras dos olas
- Desestimación de variables del proceso de encuestación irrelevantes para el modelado
- Análisis estadístico descriptivo de las escalas totales para seleccionar las features relevantes:
 - Variables categóricas:
 - Distribución de clases para comprobar si las respuestas codificadas son incorrectas
 - Evaluar si la variable es lo suficientemente interesante como posible predictor
 - Evaluar si es posible convertir la variable, con posibles variables similares, en un índice
 - Evaluar la conveniencia de aplicar target_encoder en el preprocesamiento del modelo
 - Variables numéricas:
 - Valores nulos
 - Valores cero
 - Valores lógicamente inconsistentes
 - Valores atípicos
 - Valores duplicados
- Selección de features para la fase de modelado

Este análisis fue un proceso muy laborioso y repetitivo tal y como se puede apreciar en el notebook referido, pero permitió conocer en profundidad el dataset de manera intrínseca, especialmente las diferentes escalas psicológicas que se incluyeron para caracterizar a los participantes.

Modelado

Los pasos de creación y optimización del modelo están contenidos en el notebook 02_datamodeling. Las tareas de diseño del modelo se ejecutaron de manera iterativa, hasta dar con la versión final que se empleará en la aplicación web. En concreto se realizaron las tareas siguientes:

- Recodificación del target:
 - Inicialmente podíamos encontrar cuatro clases para el estado de resiliencia en la segunda ola: sólo consideraban a una persona como “resiliente” si había mostrado ausencia de distrés tanto en la primera y como en la segunda ola. Si la persona había mostrado distrés en la primera ola pero no en la segunda, la clase asignada era “recuperado”. Se decidió agrupar ambas clases en una única clase y considerar esa unión como la clase 1 del target del modelo.
 - De igual manera, se agruparon las otras dos clases (distrés sostenido y distrés retrasado), sin tener en cuenta el distrés mostrado en la primera ola, en una sola clase “no resiliente” dando lugar a la clase 0 del target.
 - La existencia de cuatro clases abre posibilidades para seguir ampliando el alcance del modelo que se discutirán en el último punto de este documento, las conclusiones.
- Identificación de variables que utilizadas directamente en la definición empírica del target:
 - Dado que la creación del target utiliza escalas de la propia encuesta, tras entender bien cómo se construyen las categorías mencionadas anteriormente, se desestiman las siguientes features: 'PHQ_15_TOTAL', 'PHQ_9_TOTAL', 'GAD_TOTAL', 'PTSD_REEXP', 'PTSD_AVOID', 'PTSD_THREAT', 'PTSD_SEVERITY', 'PTSD_IMPAIR', 'PTSD_DIAGNOSIS', 'PTSD_TOTAL', 'SUBST_STRESS', 'DEP_DIAGN'.
- Preprocesamiento de las features:
 - Se utiliza el método TargetEncoder para tratar las features categóricas. Escogemos este método ya que el número de categorías en muchas de las features es alto y el método alternativo OneHotEncoder daría lugar a un set desproporcionado de variables respecto al número original, entorpeciendo el objetivo de encontrar qué aspectos característicos de la persona tienen mayor poder predictivo, por el aumento de la granularidad que sucedería y la dilución del efecto de las escalas y variables originales.
 - Se utiliza el método StandardScaler para tratar las features numéricas, dado que uno de los modelos que se quiere testar (LogisticRegression) depende de la escala de las variables numéricas. También queremos jugar con la aplicación de PCA en otra fase del preprocesado y de esta manera nos aseguramos de que no habrá ningún problema.
 - Se desestima el uso de PCA ya que no mejora los scores de los modelos testados y dificulta la interpretación del resultado del modelo.
- Selección del modelo clasificador:
 - Se probaron versiones simples de los modelos de clasificación siguientes, encontrados en la librería de data science de Python scikit-learn:
 - LogisticRegression
 - DecisionTreeClassifier
 - RandomForestClassifier

- Tras numerosas pruebas con diferentes parámetros e iteraciones, se decide utilizar RandomForestClassifier, con un proceso de optimización de hiperparámetros basados en el método RandomizedSearchCV.
- Se desestima LogisticRegression ya que se consiguen resultados más pobres probablemente debido a la relación no lineal que existe entre las features.
- Utilizaremos DecisionTreeClassifier para mejorar la interpretación del resultado del modelo.
- Optimización de hiperparámetros:
 - Para encontrar los valores óptimos de los hiperparámetros del modelo RandomForestClassifier se utiliza el método RandomizedSearchCV, que es preferible a la alternativa GridSearchCV para aligerar el tiempo de ejecución cuando hay un alto número de permutaciones.
 - Se trabaja sobre los siguientes hiperparámetros tratando de limitar los valores posibles de manera que los resultados sobre el set de train y de test sean similares y no se produzca overfitting:
 - 'n_estimators'
 - 'min_samples_leaf'
 - 'min_samples_split'
 - 'max_depth'
 - 'max_features'
 - Se elige optimizar sobre la métrica accuracy ya que tenemos un problema balanceado: clase 1 del target representa el 63% de las observaciones.
 - Se ejecuta sobre un fold de 3 de cross validation para disminuir el bias.
- Empaquetado para reutilización del modelo para webapp usando librería Pickle.

En este estudio se han creado tres modelos diferentes:

- Modelo de predicción de resiliencia teniendo en cuenta las variables que tenían relación directa con el contexto de pandemia COVID-19.
- Modelo de predicción de resiliencia teniendo en cuenta sólo variables caracterizadoras del participante, sin incluir el contexto directo de la pandemia.
 - La motivación para crear esta versión radicaba en tratar de aislar la capacidad predictiva del modelo del contexto concreto en el que se está caracterizando a la persona. De esta manera podemos trabajar sobre la hipótesis de que el modelo será útil para cualquier contexto difícil que esté atravesando la persona sobre la que se quiere realizar la predicción.
- Modelo simplificado de predicción de resiliencia para webapp.
 - Para poder generar una aplicación con una experiencia de usuario era necesario que la recogida de información del usuario fuera menor a 10 minutos. Por este motivo, se decidió generar esta versión utilizando las 8 features de mayor poder predictivo encontradas en la versión anterior del modelo.

La evaluación de los modelos se discute en el punto siguiente del documento.

Interpretación

El último paso del estudio llevado a cabo consiste en aplicar métodos que permitan interpretar el resultado de predicción obtenido para una observación dada. Esta interpretación servirá, además de para entender qué está pasando dentro del modelo y suavizar así el efecto black-box habitual, para dar feedback al usuario de la aplicación.

El feedback debe ser accionable, es decir, debe permitir al usuario generar una acción a partir de él. Con esta premisa, tras una búsqueda de posibilidades en la red, se encontró una librería sencilla que servía de manera ideal para el propósito. Se trata de la librería `treeinterpreter` que se puede encontrar en <https://github.com/andos/treeinterpreter>. Esta librería permite descomponer la predicción de cada clase en el bias (el valor a priori de la predicción del set de entrenamiento) más la contribución positiva o negativa de cada feature. La explicación más detallada de la metodología puede encontrarse en la entrada <https://blog.datadive.net/interpreting-random-forests/>.

Se decide utilizar una representación en gráfico en cascada para visualizar el resultado final que ofrece esta librería. En el notebook 03_interpretacion puede encontrarse el trabajo desarrollado que luego se implementará en el código de la aplicación web.

Desarrollo de webapp

Para implementar la versión simplificada del modelo, se decide utilizar `streamlit` ya que permite crear de manera sencilla con Python una aplicación web dinámica. Tiene implementado de manera nativa la conexión con github, con lo que no hace falta salir de nuestro entorno de trabajo.

Dentro de la carpeta `/resilience_webapp` en el repositorio puede encontrarse el código principal, el fichero de requerimientos para `streamlit` y el modelo empaquetado por medio de la librería `Pickle`.

El trabajo desarrollado para implementar el modelo consistió en tipificar los cuestionarios de las escalas escogidas, crear el código de recodificación y cálculo de la métrica total para cada escala, crear el vector final para que, tras cargar el modelo podamos dar lugar a la predicción y al feedback descrito en el apartado anterior.

La aplicación es accesible en la url:

<https://ignaciojdelrey-resili-resilience-webappresilience-webapp-iu553f.streamlitapp.com/>

4. Resumen de resultados

Se ha evaluado la eficacia de los modelos según la mejora sobre el prior conseguida a la hora de clasificar como resiliente una observación y analizando si existe o no overfitting en el entrenamiento. Para esta evaluación utilizamos la métrica accuracy. También se han calculado las métricas precision y recall, así como f1_score para tener mejor perspectiva el desempeño de los modelos. Por otro lado, con la intención de comprobar que el modelo no estuviera haciendo overfitting, se graficaron las curvas roc y se calcularon las áreas correspondientes.

El resumen de resultados se encuentra en la siguiente tabla:

Modelo	Accuracy		Precision		Recall		F1_score		ROC_area	
	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
RF_Covid19	0,80	0,76	0,80	0,77	0,90	0,89	0,85	0,83	0,76	0,70
RF_general	0,78	0,75	0,77	0,76	0,91	0,90	0,84	0,82	0,73	0,68
RF_simple	0,77	0,75	0,77	0,76	0,90	0,89	0,83	0,82	0,72	0,68

Los resultados son evaluados como muy buenos, ya que presentan valores altos de capacidad predictiva de la clase 1 (ser resiliente) sin un incremento excesivo de los falsos positivos, como se puede apreciar en el valor F1_score conseguido. El valor de accuracy obtenido en el modelo más sencillo es un 22% mayor que el prior (0,63) usando un baseline ZeroR.

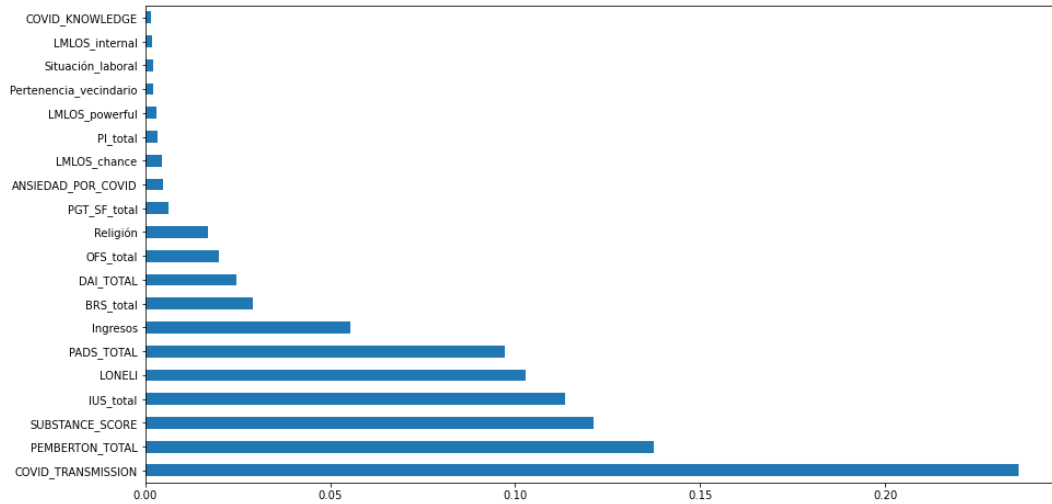
Además, se puede observar que los resultados son bastante similares entre los modelos, lo cual implica que existe mucha información compartida entre las features de entrenamiento, que no se pierde al descartar variables al crear versiones más sencillas del modelo inicial.

Por otro lado, se ha analizado la importancia de las features para predecir la resiliencia en cada modelo, encontrándose consistencia, con algunos matices relevantes:

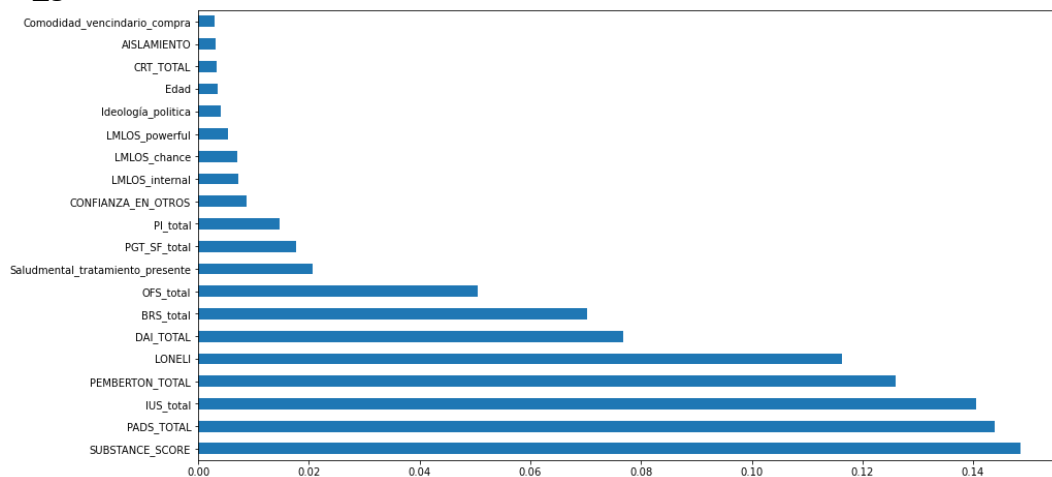
- El mayor poder discriminatorio en el modelo general proviene del conocimiento específico sobre los modos de contagio del Covid-19. En las primeras iteraciones de modelado, probando el modelo DecisionTreeClassifier, se pudo analizar cómo esta variable discriminaba, y lo hacía de manera que “saber mucho del Covid” te hacía menos resiliente y viceversa.
- La importancia de variables caracterizadoras sociodemográficas es baja y desaparece cuando se deja de tener en cuenta el contexto concreto ante el que se mide la resiliencia (variables Covid_19). La variable que con mayor constancia ha aparecido en las iteraciones de modelado ha sido ‘Ingresos’, indicando que un menor nivel de ingresos estaba asociado a menor capacidad de resiliencia.
- Las variables caracterizadoras psicológicas están entre las de mayor poder predictivo en todos los modelos, y de hecho son más importantes sin contexto. Las más relevantes son, por un lado las escalas que miden el bienestar personal (‘PEMBERTON_TOTAL’), la apertura al futuro (‘OFS_total’) y la capacidad auto-percibida de afrontar dificultades (‘BRS_total’), todas ellas relacionando valores altos en la escala con mayor capacidad de resiliencia; y por otro lado, la intolerancia a la incertidumbre (‘IUS_TOTAL’), pensamientos paranoicos (‘PADS_TOTAL’) y la ansiedad ante la muerte (‘DAI_TOTAL’), todas ellas relacionando valores altos en las escalas con menor capacidad de resiliencia.

- La variable de relaciones interpersonales 'LONELI' aparece también con alto poder discriminante en todos los modelos, así como la de 'SUBSTANCE_SCORE'. En ambas, valores más altos se asocian a menor capacidad de resiliencia y viceversa, indicando la importancia del soporte social y los hábitos de vida saludable en la capacidad de resiliencia.

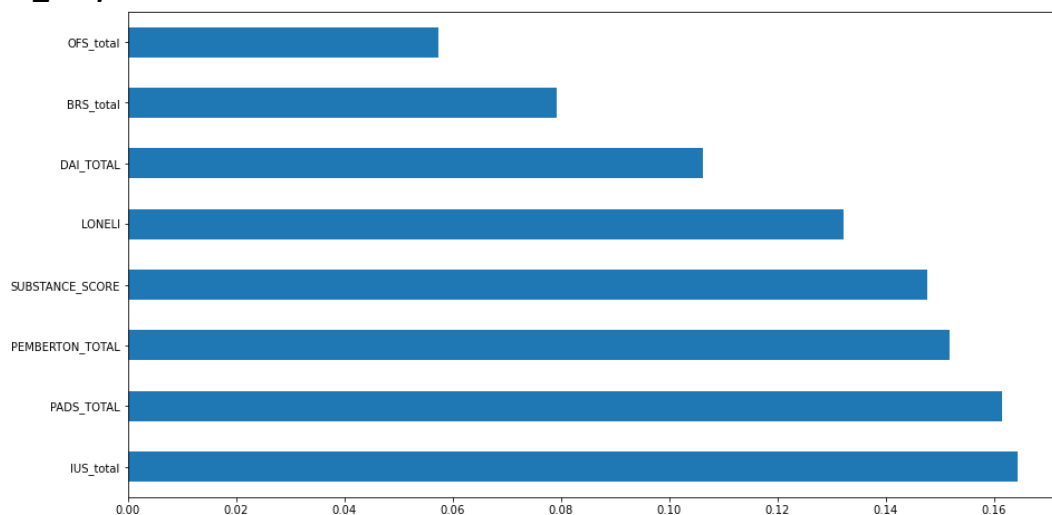
RF_Covid19



RF_general



RF_simple



5. Conclusiones

El resultado del estudio es satisfactorio, ya que se han cumplido los objetivos planteados inicialmente:

- Se ha conseguido desarrollar un modelo exitoso para predecir la capacidad de resiliencia de una persona en un futuro próximo a partir de una caracterización sociodemográfica, psicológica y del contexto que afronta.
- Se ha demostrado que el efecto del contexto está implícito las medidas psicológicas utilizadas para caracterizar a la persona, ya que la eficacia de los modelos no disminuye al eliminar las variables que se refieren explícitamente a cómo está viviendo o cómo afronta la persona la situación ante la que se probará su capacidad de resiliencia.
- Se ha validado parcialmente los resultados obtenidos en el estudio original del que partía este trabajo. La diferencia fundamental reside en que en nuestra investigación no hemos visto que las variables caracterizadoras sociodemográficas tengan poder predictivo para la capacidad de resiliencia de la persona, sino que más bien parece que este poder se consigue con éxito utilizando tan sólo ciertas escalas de caracterización psicológica discutidas en el punto anterior, algunas de las cuáles sí coinciden con las encontradas en el estudio del equipo de la UCM.
- Se ha desarrollado una versión del modelo implementable en una aplicación interactiva con buena experiencia de usuario, sin merma de la eficacia, y con capacidad de ofrecer feedback accionable al usuario para trabajar en factores que mejoren su capacidad de resiliencia.

6. Manual de usuario para la aplicación interactiva

La aplicación es accesible en la url:

<https://ignaciojdelrey-resili-resilience-webappresilience-webapp-iu553f.streamlitapp.com/>

Seguir las instrucciones mostradas al cargar la aplicación.