Trabajo de Fin de Máster PEC 3: Desarrollo del trabajo. Fase 2.

Pablo Núñez de Arenas Martínez

29/5/2023

**Análisis y selección del modelo:**

A continuación, se muestra una tabla resumen con los resultados de cada modelo.

Tabla

Descripción generada automáticamente

Tabla 1: Resumen de todos los modelos.

Entre todos los modelos, se podían destacar el SVM lineal con 7 variables, el KNN con 3 variables o el random forest (40) con 3 variables. Finalmente, se creyó que la mejor opción era el KNN con 3 variables. El SVM lineal no convencía del todo, ya que algunas de las variables seleccionadas estaban sesgadas. En cuanto al random forest, resultaba extraño que se requirieran tantos árboles para obtener un modelo óptimo utilizando solo 3 variables.

En conclusión, se optó por el modelo KNN con 3 variables debido a que parecían ser las mejores opciones y porque es un modelo fácil de implementar y adecuado para conjuntos de datos con pocas muestras o variables, como era el caso.

**Definición del esquema de la aplicación:**

La idea inicial consistía en desarrollar una aplicación en la que el paciente ingresara los datos de las variables seleccionadas, y luego el modelo le proporcionaría la predicción correspondiente, la cual se mostraría en pantalla.

Vamos a mantener esta idea, pero realizaremos algunos ajustes. Además de crear una página para realizar predicciones, intentaremos agregar una segunda página que permita a los usuarios explorar la base de datos utilizada para crear el modelo predictivo.

En lugar de que el paciente introduzca manualmente los valores de las variables, se le ofrecerá la opción de seleccionar entre diferentes opciones. De esta manera, se evitarán errores causados por la introducción de valores negativos, fuera de rango o no contemplados por el modelo.

Para que el paciente pueda ubicar cada variable en su respectivo grupo, se mostrará en pantalla el rango de valores correspondiente a cada nivel de la variable.

A continuación, se muestra un croquis del diseño inicial de la aplicación:

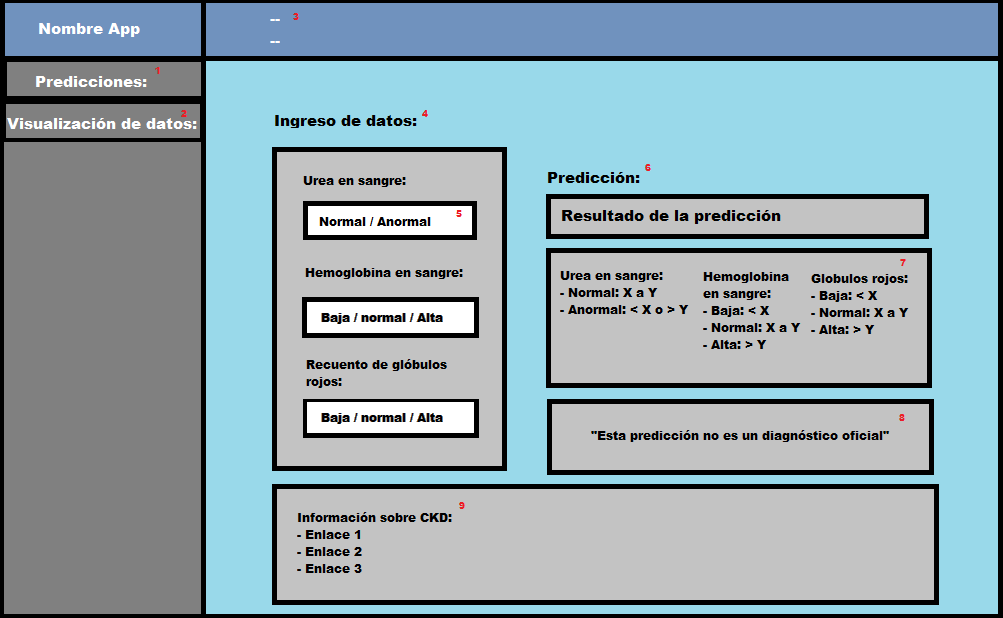


Ilustración : Croquis de la APP mostrando la página “Predicciones”.

1. Es la ventana “Predicciones” si hacemos click en esta pestaña nos aparecerá la página mostrada en el croquis donde deberemos seleccionar los niveles de la variable para recibir la predicción.
2. Nos lleva a la parte de la App que nos permitirá explorar la base de datos utilizada para el modelo.
3. Este botón nos permitirá ampliar o reducir el tamaño de la página quitando o poniendo las ventanas de predicción o visualización (ventana gris).
4. Parte de la pantalla en la que seleccionaremos los niveles de las variables.
5. Cada variable tendrá un cartel desplegable en el que podremos seleccionar los niveles.
6. En esta parte de la pantalla recibiremos los resultados.
7. Este cuadro será el esquema que permitirá al paciente saber a que nivel pertenece en cada variable.
8. Se añadirá un cuadro que informe al paciente de que esta predicción no constituye un diagnóstico alentándole a acudir al médico en caso de que sea necesario.
9. Se les aportará a los pacientes información extra sobre la enfermedad.

La segunda de las páginas se pensó en añadirla para ampliar el trabajo y ofrecer más opciones de visualización de datos. Esta página adicional no tiene una utilidad real para los pacientes en términos de diagnóstico o tratamiento de enfermedades renales. Su propósito principal es mostrar la relación entre diferentes variables y permitir una exploración más detallada de los datos utilizados en el modelo predictivo.

La página de “visualización de datos” tendrá la siguiente estructura:

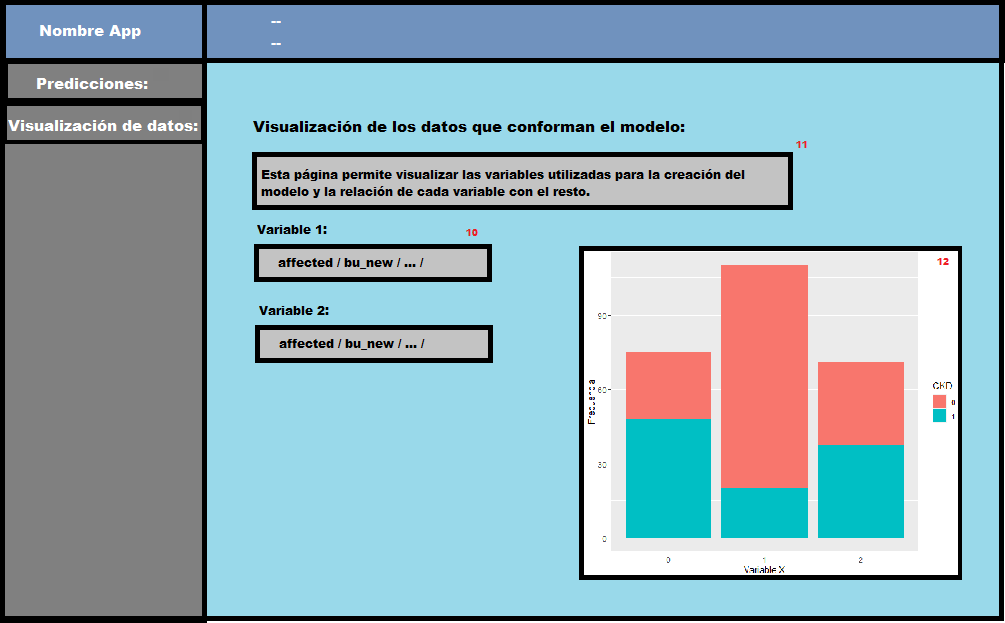


Ilustración : Croquis de la APP mostrando la página “Visualización de datos”.

1. En esta sección habrá dos botones, el primero permitirá escoger el número de columnas que aparezcan en el histograma, es decir si se escoge una variable con 2 niveles aparecerán dos columnas, si se escoge una con 3 aparecerán 3. El segundo botón permitirá controlar los grupos en que se divide cada columna.

Por ejemplo, el histograma que aparece en el croquis se genera al escoger en primer lugar la variable “hemo\_new” y después “affected”. Veríamos la proporción de casos de CKD en cada nivel de la variable “hemo\_new”.

1. Este cartel simplemente sirve para explicar el funcionamiento de la página.
2. Resultado de la visualización.

**Desarrollo y comprobación del funcionamiento:**

Hemos desarrollado y comprobado el funcionamiento de la APP utilizando Shinydashboard. Esta librería de R permite usar el lenguaje HTML para definir la estructura de la aplicación y los diferentes elementos visuales, como encabezados, paneles, botones y tablas. Después se usa CSS para personalizar el estilo, el diseño y los colores de los elementos HTML.

La estructura básica de una APP en Shinydashboard consta de:

1. Interfaz de usuario (ui): define la apariencia y la estructura de la interfaz de usuario de la aplicación. Podemos usar estructuras como *dashboardHeader,* *dashboardSidebar, dashboardBody* o *tabItems* etc*.* para añadir encabezados, barras laterales, cuerpos o secciones de la aplicación.
2. Server: Esta sección contiene el código que define la lógica y el comportamiento de la aplicación. En el servidor, se definen los componentes de salida (output) que representan los resultados o visualizaciones que se mostrarán en la interfaz de usuario. También se pueden utilizar los componentes de entrada (input) para capturar las selecciones o acciones del usuario.
3. shinyApp(ui, server): La función shinyApp se utiliza para combinar la interfaz de usuario (ui) y el servidor (server) en una aplicación Shiny completa.

En una aplicación Shinydashboard, el código se escribe en archivos R con extensión ".R" o ".Rmd" , donde se combinan instrucciones R, HTML y CSS de manera integrada. Estos archivos contienen la lógica de la aplicación, incluyendo la definición de la interfaz gráfica, la carga de datos, la implementación de funciones y la generación de resultados dinámicos.

Respecto al desarrollo de la aplicación en sí, esta se realizó con bastante facilidad e incluso antes de las fechas contempladas inicialmente lo que nos permitió agregar una segunda parte a la APP que permitiese visualizar las variables utilizadas para la creación del modelo.

Esta adición a la APP nos dio algunos problemas a la hora de mantener una correcta identación y legibilidad del código, al final a base de prueba y error se consiguió desarrollar la APP con un código funcional y bastante bien optimizado respecto a versiones anteriores.