# TP 2: Diseño de la solución

En el presente trabajo práctico se busca avanzar en busca de la solución al Business Case presentado y explorado en el TP1.

Se espera que surja de esta etapa una solución que involucre una parte tecnológica en la cual se incorpore al menos un modelo de Machine Learning. Es importante destacar que la solución final al problema puede no ser 100% tecnológica, por ejemplo, involucrando personas utilizando una aplicación o página y realizando acciones que en su conjunto mueven los KPIs planteados durante la presentación del Business Case en el TP1.

El trabajo sigue siendo con los mismos **grupos del TP1 a menos que se autorice lo contrario por los docentes**.

La solución a este TP se deberá entregar **impreso en papel y subir al campus**. En el documento se espera encontrar:

1. Diseño general de la solución
   1. Uno o dos wireframes que ayuden a visualizar la solución
   2. Descripción conceptual de la solución argumentando como mueve el KPI (máximo 1/2 carilla)
2. Identificación de la parte que involucra Machine Learning
   1. Familias de técnicas a usar (clasificación, regresión, etc)
   2. Descripción conceptual sobre cómo el modelo se integra en la solución final (máximo 1/2 carilla)

## 0. Business Case

Una clínica quiere reducir las muertes de sus pacientes por ACV: se requiere de la

identificación de los pacientes de alto riesgo para poder brindarles, a sus respectivos

médicos de cabecera, una notificación temprana del caso en pos de tomar las medidas

de prevención necesarias.

Nuestro KPI será la reducción de las muertes por accidentes cardiovasculares anuales en un 20% en un lapso de 5 años en nuestro hospital.

## 1. Diseño general de la solución

La solución a nuestro problema es la implementación de un software predictivo vinculado con la base de datos de cada paciente en pos de identificar el grado de riesgo que tienen los pacientes de padecer un ACV. Esto permitiría luego en conjunto con el médico (quien posee en este caso conocimiento del dominio) recomendar un tratamiento y las precauciones necesarias para prevenir el ACV.

A continuación presentaremos los wireframes de nuestra solución



Fig.1. Wireframe para dispositivo móvil

Lo que observamos en la figura 1 es la versión móvil de la aplicación. Está versión únicamente estaría presente de manera secundaria (como versión lite y portátil) para los médicos. En donde se ven tanto características y diversas métricas de cada paciente como “el buscador” de pacientes.

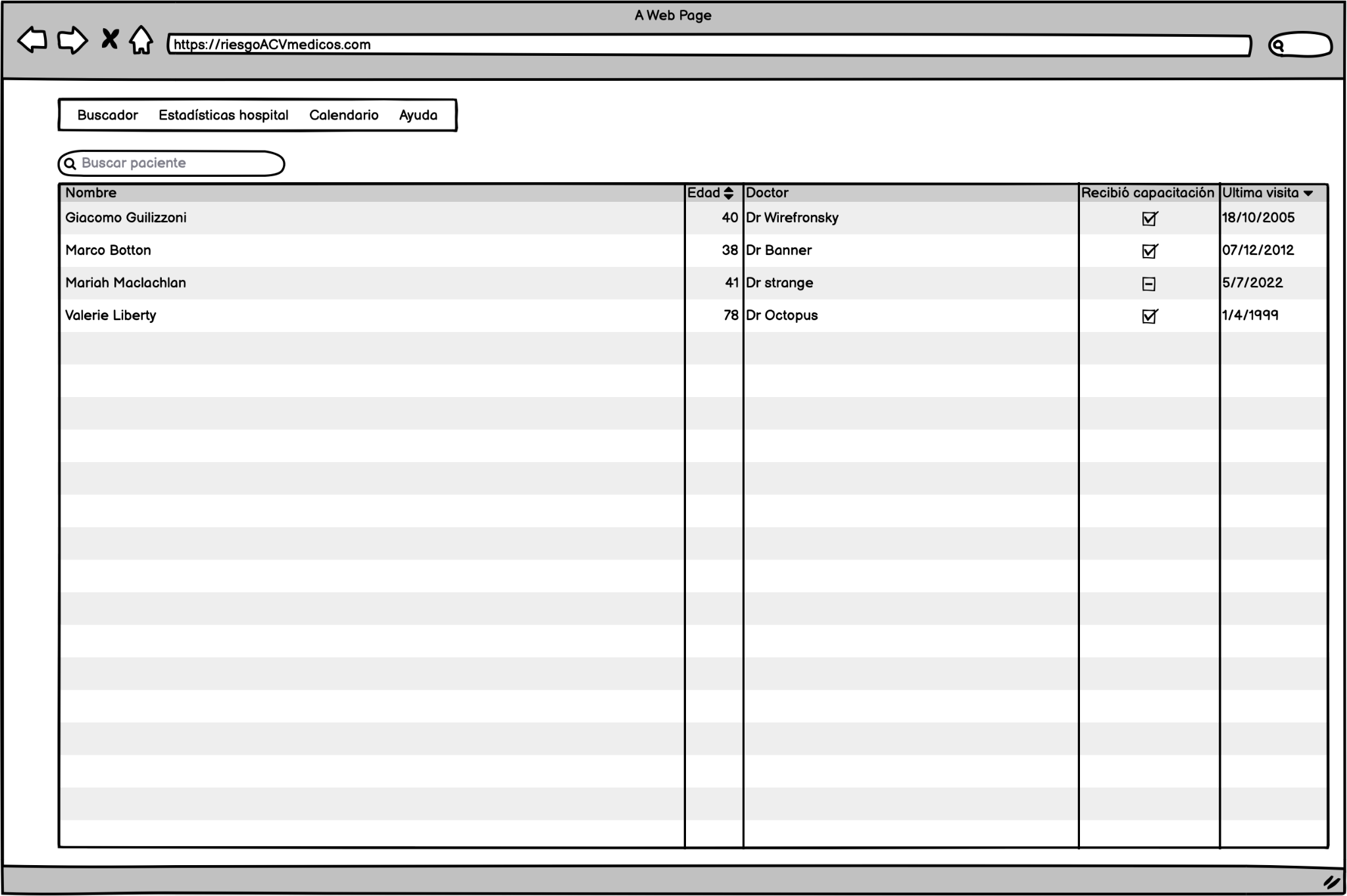


Fig.2. Wireframe menu principal version escritorio. l

En la figura 2 tenemos las diversas opciones que nos ofrece la aplicación, desde buscador de paciente, manejo de turnos y estadísticas generales del hospital que sirvan de seguimiento de la inferencia que tienen los médicos a nivel más global.

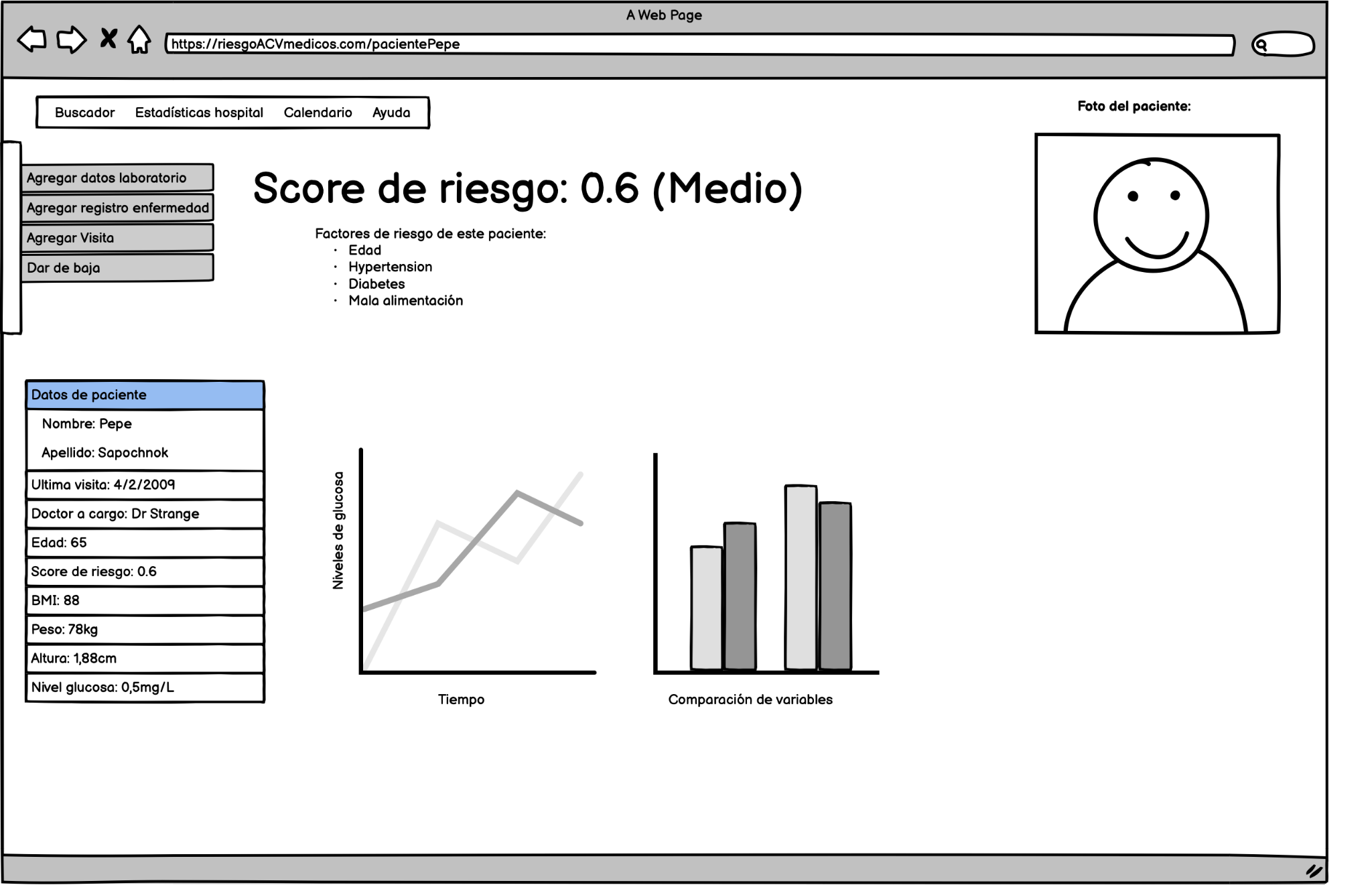


Fig.3. Wireframe página de un paciente (versión escritorio).

El wireframe mostrado en la figura 3 responde las siguientes preguntas: ¿qué podemos observar en cada paciente?, ¿cómo evolucionaron sus estudios médicos en el tiempo?, su última visita, y la comparación de algunas variables en relación al grupo etario o género que pertenece el paciente. Nótese que el score de riesgo no solamente nos dice su nivel sino que además el modelo nos justifica qué es lo que hace que el modelo tenga el score que tiene. Está información ponerla a disposición del médico puede en algunos casos serle útil para tomar decisiones (si así lo desea el médico).

## 2. Identificación de la parte que involucra Machine Learning

Para resolver nuestro problema, consideramos pertinente un modelo de regresión, pues al tomar el valor de la regresión entre (0,1) obtenemos nuestro score de riesgo de ACV. Asimismo, cabe destacar la importancia de la interpretabilidad de nuestro modelo por cuestiones éticas; dado que estamos tratando con pacientes, las decisiones del modelo tendrían grandes implicancias en la vida de ellos.

**Integración del modelo en la solución final**

El modelo deberá ser entrenado en una primera etapa con datos offline que ya se encuentren etiquetados. De esta manera obtendrá sus capacidades predictivas para que pueda establecerse en producción. Luego, su efectividad puede mejorarse a medida que vayan surgiendo nuevos datos, en un proceso conocido como transferencia de aprendizaje. Las predicciones sobre el estado de riesgo del paciente, es decir, el score, se interpretan en base a una escala que establecimos, y que consta de tres niveles a priori: bajo riesgo (<0.5), medio (0.5 - 0.8) y alto (0.8<).

Suponemos que la inferencia de los médicos es acertada y que los exámenes de salud de los pacientes se realizan con frecuencia y continuidad. Por lo tanto, se espera una reducción del score de los pacientes de alto riesgo a niveles normales al ser identificados y tratados. Si bien es cierto que el éxito del proyecto implica que nuestro modelo deje de predecir los niveles de alto riesgo de ACV, asumimos que si la disminución en el número de pacientes de riesgo predichos por el modelo se produce a la par de una reducción estadística de los pacientes del hospital que contraen ACV, entonces nuestro modelo habrá tenido un efecto positivo en los esfuerzos de prevención de la enfermedad. Para esto debemos tener suficientes datos antes y después de la implementación de nuestro modelo en producción. En otras palabras, para evaluar el desempeño de nuestro modelo y vincularlo con el KPI que propusimos, observamos que a corto plazo, se espera que los pacientes con un score elevado mejoren su estado de salud tras seguir las recomendaciones de su médico. Esto debe reflejarse en un descenso en el nivel de dicho score. Por otro lado, se espera que a largo plazo las estadísticas de muertes por ACV disminuyan significativamente.