El proyecto consta de un cuerpo donde se toma el input del usuario y de 3 partes principales:

* Preprocesamiento: Aquí utilizamos distintos métodos para generar una salida de datos sin errores de ortografía, letras repetidas y corrigiendo algunas omisiones de letras o incluso si hay letras en el orden incorrecto en una palabra. Pero como se puede imaginar, esto es un proceso lento y computacionalmente muy caro, por lo que a fines prácticos es mas provechoso evitarlo y esperar un correcto uso de las palabras por parte del usuario.
* Búsqueda de la respuesta: Para poder determinar que respuesta darle al usuario se utilizan 3 inteligencias artificiales independientes que determinan la carrera a la que es destinada la pregunta, la intención y subintencion del usuario y el w5 (que, como, cuando, donde, porque)
* Por último, y luego de darle la respuesta al usuario, la misma se almacena en una base de datos mongoDB junto con la pregunta, la pregunta procesada y las respuestas de las IA, junto con una nueva variable que determina si la pregunta no fue encontrada para así poder determinar en un futuro cuales son los fallos principales del sistema y así mejorarlo

Para realizar todo esto, se utilizan distintas librerías tales como

* Pymongo: que nos genera la conectividad entre el programa y la base de datos en la nube.
* Pandas: solo se utiliza por comodidad al momento de ver los documentos y en un modelo de producción se podría evitar para ahorrar espacio y poder de cómputo, agilizando así la tarea del chat.
* Regex y spacy: Ambas son utilizadas para procesar el lenguaje y generar datos que pueda ser suministrado a una IA sin más problemas.
* Numpy y sklern: estos son utilizados para los modelos predictivos y el proceso de los DF que utilizan
* Pickle: este mismo es el que almacena los modelos entrenados para evitar pasar por la espera de entrenarlos en cada interacción con el chatbot

Como una reflexión final podríamos observar algunos detalles tales como el hecho de que la información es escasa, los modelos son bastante ineficientes y esto lleva a que el producto final también lo sea. Para este caso la mejor solución hubiese sido posiblemente utilizar solo un sistema de matcheo entre las palabras estematizadas y la información que teníamos. Pero se opto por mantener el sistema predictivo por el hecho que dicha información en una empresa que busque hacer un bot para evitar el uso de humanos estaría mas dispuesta a brindar grandes cantidades de información que ayudarían al proceso.

También podemos decir que el sistema es bastante certero y si se evita el procesamiento de corrección de los datos, seria prácticamente en tiempo real, lo que se ajusta a las necesidades de un posible cliente. Junto con el hecho que se va almacenando la información en una base de datos, la misma serviría para, en poco tiempo, mejorarla hasta el nivel deseado al poder anexar casos reales a la información de la IA.