Introducción a la privacidad diferencial

Es un marco que se define sobre operaciones que uno va a hacer sobre una base de datos en la que hay cuasi identificadores y datos privados que se quieren proteger. Cada usuario sería una fila en dicha base de datos. Tenemos un curador que sigue el protocolo y tiene la confianza de los usuarios que publica los datos utilizando un mecanismo M que tiene una salida R=M(D). El curador calcula funciones de los datos y tenemos un analista que trata de hacer inferencias de lo que hay en la base de datos a partir de lo que ve. A través del estudio, el analista puede encontrar una correlación entre dos elementos de la base de datos, por ejemplo, si fuera una base de datos médicos, si tiene un campo de si alguien fuma y otra de si tiene cancer, puede encontrar la relación de que si alguien fuma, entonces tiene altas probabilidades de tener también cáncer.

La privacidad diferencial protege contra el aprendizaje de cosas que no se pueden obtener mediante inferencia de la información lateral. Si tenemos dos bases de datos y en una tenemos a cierto sujeto y este no está en la otra, la privacidad diferencial busca estadísticamente que el resultado de respuesta de ambas bases de datos cuando se hace una solicitud, el resultado no sea distinguible. Se trata de que cuando el curador responda meta ruido en la respuesta para que ambas bases de datos no sean distinguibles. Lo malo de la privacidad diferencial es que puede inutilizar las bases de datos al introducir ruido en exceso. La respuesta del curador debe ser determinista (Que tenga una entrada aleatoria).

La idea es responder siempre lo más parecido independientemente de quien esté en las bases de datos. La privacidad diferencial se define sobre bases de datos vecinas, que son bases de datos muy similares en las que la diferencia es una fila que ha sido eliminado o que cambia con respecto a otra.

Como definir distribuciones similares

Dos distribuciones son parecidas si al calcular la diferencia entre una y otra no se pasa de un valor P que es el indicador de como de privadas son.