

Directrices de anotación

Antecedentes:

Estamos creando mejores formas de probar los modelos para la detección automática del discurso de odio en línea. Para ello, hemos recopilado un conjunto de datos compuesto por alrededor de mil entradas de texto breves. Cada entrada refleja un aspecto particular del odio en línea.

Sus tareas:

1. Indicar si cada entrada asignada es de **odio** o **no-odio**.
2. Marcar cualquier entrada que considere **no realista**.

¿Qué entendemos por "odio"?

Definimos el **odio** como **el abuso dirigido a un grupo protegido o a sus miembros por formar parte de ese grupo**. Los grupos protegidos se conforman según la edad, la discapacidad, la raza (color, nacionalidad, orígenes étnicos o nacionales), la religión o las creencias, el sexo y la orientación sexual, así como la identidad de género.

¿Qué entendemos por "no realista"?

Esto lo dejamos deliberadamente a su criterio. Construimos todas las entradas con la intención de que sean realistas, por lo que este indicador "no realista" es opcional y solo pretende garantizar la calidad de los datos.

Algunas cosas a tener en cuenta al anotar:

- **El lenguaje de odio en algunos contextos puede tener usos que no lo sean** (por ejemplo, el discurso contrario que hace referencia a insultos de odio: "No está bien decirle a alguien "negro").
- **El destinatario del abuso es importante para determinar si algo es discurso de odio o no**. El abuso interpersonal y el abuso contra grupos no protegidos, como profesiones y afiliaciones, no es discurso de odio (por ejemplo, "odio a los judíos" es discurso de odio, pero "te odio", "odio mi mesa" y "odio a los médicos" no lo es).

Otras orientaciones clave:

1. No marque las entradas como no realistas solo porque crea poco probable que puedan aparecer en Internet. Marque solo las entradas sin sentido o gramaticalmente incorrectas.
2. Por favor complete sus anotaciones de forma independiente y no las comente con otras personas.
3. La mayoría de los enunciados son bastante cortos, así que no los piense demasiado.

¡Muchas gracias por su trabajo!