DESCRIPCIÓN y CONCLUSIONES

archivo “Formulario\_clasificación muestras ELE\_respuestas.xlsx”:

(hoja: original data)

Conclusiones de los expertos:

- les resulta difícil al no tener contexto, alguno se lo ha imaginado. No pueden valorar si un descuido es ocasional o una carencia

- entienden como problema que la realidad es que un estudiante de cualquier nivel, puede realizar una frase perfecta o completa de errores dependiendo el momento

- una frase aislada puede ser dicha correctamente por hablantes de distintos niveles

- los fallos pueden ser puntuales: "El hecho de que escriba "Estimdos" en vez de "Estimados" no indica que sea de un nivel más bajo."

- en ciertos casos han valorado de acuerdo al vocabulario

(hoja: table)

En esta hoja están los resultados de los expertos y del caes, y se han aplicado fórmulas sobre ellos para obtener mayor información de los datos.

(hoja: table2)

En esta hoja se muestra la relación que tienen los resultados de los expertos en relación al nivel atribuido por CAES.

La columna:

- aciertos, trata del número de coincidencias que han tenido los expertos en cada frase con el nivel del CAES

- niveles por debajo, indica de las tres evaluaciones hechas por los expertos cuántas corresponden con un nivel inferior al del CAES

- niveles por arriba, indica de las tres evaluaciones hechas por los expertos cuántas corresponden con un nivel superior al del CAES

- coincidencias humanas, indica cuantas de las evaluaciones de los expertos coinciden entre sí (no tienen por qué coincidir con el CAES)

- nivel humano mayoritario, indica cual es el nivel en el que más han coincidido los expertos (no tiene por qué coincidir con el CAES)

De aquí se obtiene que:

- sólo hay una frase en la que coinciden los tres expertos y el CAES

- Hay 3 frases en las que dos expertos coinciden con el CAES

- Hay 18 frases en las que un experto coincide con el CAES (los otros dos no)

- Hay 18 frases en las que ningún experto coincide con el CAES, ni los expertos coinciden entre ellos

(hoja: heatmap)

Muestra con colores como son las evaluaciones de los expertos en relación con la clasificación del CAES.

(hoja: caes vs expertos)

**Tabla 1:** muestra el número de veces que los expertos han evaluado una determinada frase del CAES en un nivel.

Lo ideal sería que la diagonal tuviese números más altos que el resto, pero como se puede ver esto no ocurre.

**Tabla 2:** muestra con colores la misma información que la Tabla 1, la idea es ver la tendencia de degradado. Se puede observar que el triángulo inferior de la diagonal y el superior tienden a clarear. Esto quiere decir, que el nivel más bajo, pocas veces es evaluado como un nivel alto, y viceversa, un nivel alto, pocas veces es evaluado como un nivel bajo.

**Tabla 3:** (no sé si tiene sentido) muestra la relación de los niveles del CAES con el nivel en que al menos dos expertos estaban de acuerdo. Se puede observar que la diagonal está vacía (casi) y esto muestra que no hay casi coincidencias entre lo que evalúan los expertos y el CAES, y las que hay se encuentran en el nivel más bajo (A1).

(hoja: expertos fallos)

La primera tabla está ordenada por nivel de coincidencia entre humanos, el color verde marca cuando un experto coincide con el nivel del CAES.

La segunda tabla está ordenada por nivel del CAES, la idea era ver si hay algún nivel que le cuesta más a los expertos que otro, o que disienten más entre ellos. No veo una conclusión clara al respecto.

(hoja: model pred.)

Son los datos que devuelve el modelo.

(hoja: model pred fallos)

- hay 31 aciertos y 9 fallos

Las frases están ordenadas por coincidencias, y están marcadas de naranja las frases que no coinciden.

Una conclusión podría ser que el modelo se equivoca en frases cortas (como "buenos días" o abrazos"), pero en los aciertos también hay frases cortas.

(hoja: model pred. vs caes vs expr)

**Tabla 1:** muestra la relación de las predicciones del modelo con CAES, se observa que la mayoría de los datos se encuentran en la diagonal principal, y que las predicciones que no son correctas se encuentran próximas al nivel que marca el CAES.

**Tabla 2:** (creo que no tiene sentido) muestra la relación de las predicciones del modelo con las de los expertos.

(hoja: model vs caes vs expr II)

Muestra con colores, la relación entre los niveles dichos por los expertos, los del modelo y los del CAES.

De los 9 que no acierta el modelo con el CAES llama la atención que en 6 coincida con al menos un experto, de hecho, en 5 coincide con al menos dos expertos en la decisión.

Una de las 9 frases que no acierta el modelo no coincide tampoco con ninguna respuesta de los expertos.

(hoja: Heatmap Completo)

Heatmap completo de las tres valoraciones

(hoja: all vs perplejidad)

Se ha calculado la media de la perplejidad de todos los datos (todos, los que se están “prediciendo” también) y la desviación estándar.

Se ha establecido como threshold entre los niveles, la media del nivel más menos la mitad de la diferencia que hay con el nivel superior e inferior.