

# Hacia un clasificador automático de emociones en texto para considerar las emociones en el proceso de aprendizaje en una institución de educación superior

Jossie Murcia<sup>1</sup>, Sebastian Moreno<sup>2</sup>, and María-Irma Díaz-Rozo<sup>3</sup>

<sup>1</sup> Escuela Colombiana de Ingeniería Julio Garavito  
sebastian.moreno-r@mail.escuelaing.edu.co

<sup>2</sup> Escuela Colombiana de Ingeniería Julio Garavito  
jossie.murcia@mail.escuelaing.edu.co

<sup>3</sup> Escuela Colombiana de Ingeniería Julio Garavito  
maria.diaz@escuelaing.edu.co

**Abstract** Several works have proposed text-based emotion classification models. However, a Spanish language approach has not been studied enough and, generally, no study considers the context and demography drawbacks, implicit in the text. Automatic classification of basic emotions considered on the International Survey on Emotion Antecedents and Reactions (ISEAR) dataset were treated (trained and validated) with an originally designed ensemble classifier schema. SemEval 2007 was used to validate our results with other investigations. The ensemble schema is based on the assumption of combining a batch of statistical machine learning classification models (NaïveBayes, Support Vector Machines and Neural networks models) and a knowledge-based tool (NRC Word-Emotion Association Lexicon), aiming to benefit from their competence to better classify certain emotions drawbacks. The deterministic voting schema behaves in a similar way to the best results of each model. The results affirm that the problem context directly affects the model results. To determine the inconveniences that may arise due to a dependency on a language and a context, two approaches were proposed; ISEAR dataset translated into Spanish and our own Academic Survey on Emotional Reactions in Computational Learning Processes (ASER-CLP). The analysis performed with ISEAR translated shows similar results compared to the best-trained model with the original dataset (an overall F-measure value of 62,9 % vs 66,4 %; respectively). The translated datasets can be useful in the same Spanish context.

**Resumen** En varios trabajos han propuesto modelos de clasificación de emociones en texto. Sin embargo, la clasificación de texto en español no ha sido ampliamente estudiada y, en general, ningún estudio considera los trasfondos del contexto y la demografía implícitos en el texto. Con el fin de apropiar los avances en el área, se desarrolló un modelo con un diseño de ensamblaje original para la clasificación automática de las emociones

básicas consideradas en el conjunto de datos de la Encuesta Internacional sobre Antecedentes y Reacciones de las Emociones (ISEAR). Se usó SemEval2007 para validar los resultados con otras investigaciones. El esquema de ensamblaje parte del supuesto que al combinar modelos de aprendizaje automático (Redes Bayesianas, Máquina de Soporte Vectorial y Redes Neuronales) y modelos basados en conocimiento (Léxico de Asociación Palabra-Emoción; NRC) se aprovecha la capacidad individual de clasificación de cada modelo. Los resultados confirman que este esquema de combinación se comporta de manera similar al modelo con los mejores resultados, y que el contexto del problema afecta directamente los resultados del clasificador. En aras de determinar los inconvenientes que pueden surgir debido a la dependencia de un idioma y un contexto, se propusieron dos acciones: traducir el conjunto de datos ISEAR a español y construir un conjunto propio de datos (ASER-CLP) con base en una encuesta académica sobre reacciones emocionales en procesos de aprendizaje computacional. El análisis, realizado con ISEAR-traducido, muestra resultados similares con el modelo mejor entrenado, utilizando el conjunto de datos original (un valor F-measure de 62,9 % y 66,4 %, respectivamente). Los conjuntos de datos obtenidos pueden ser útiles para futuras investigaciones en español.

## 1. Introducción

Las emociones juegan un papel fundamental en la vida. La existencia de cada ser humano es producto de pensamientos y conductas personales y sociales determinados por el estado afectivo de cada individuo o de la población. Así, el estudio de las emociones permite comprender las distintas formas de pensar y actuar del ser humano bajo su influencia; en particular, de estudiantes de pregrado en un entorno de aprendizaje que exige un alto rendimiento.

La inteligencia emocional (IE) propone entender la inteligencia más allá de los aspectos cognitivos. En educación las investigaciones han evidenciado que para el logro de objetivos de aprendizaje, cuando el coeficiente intelectual no es un problema, el emocional es el que marca la diferencia. El primer pilar de la inteligencia emocional es el reconocimiento de la emoción. Este es el pilar de este proyecto de investigación: identificar las emociones que expresan los estudiantes para que logren aprender a asimilarlas, comprenderlas y regularlas potenciando su proceso de aprendizaje.

La computación afectiva es un área de investigación que tiene como propósito lograr que las emociones puedan ser gestionadas automáticamente por una computadora. Esta área de investigación estudia la detección, síntesis y expresión de emociones.

Los trabajos de reconocimiento de emociones usan diferentes aproximaciones para la detección: la voz, el texto, las expresiones faciales, los gestos y el movimiento corporal [1]. No obstante, uno de los principales medios de comunicación que tienen actualmente las personas, debido al uso de tecnologías de información, son los medios textuales. A partir de la identificación de los patrones de expresión

del lenguaje de una persona se puede deducir la emoción subyacente [6], lo que indica que es posible abstraer la emoción manifestada bajo un texto.

La propuesta de éste trabajo de este trabajo es desarrollar modelo compuesto para la clasificación de emociones, a partir de la integración de técnicas de aprendizaje automática y basadas en conocimiento. Para la investigación se usaron varios conjuntos de datos: ISEAR <sup>4</sup>), SemEval y ASER-CLP (un conjunto de datos propio obtenido a partir de una encuesta académica a los estudiantes).

El presente artículo está estructurado de la siguiente forma. La sección 2 describe las teorías y trabajos relacionadas mientras la sección 3 presenta los principales aspectos de la metodología adoptada en el proyecto. La sección 4 discute los resultados obtenidos hasta la fecha, y la sección 5 expresa una discusión de las conclusiones y el trabajo futuro.

## 2. Revisión de la literatura

### 2.1. Marco referencial

Con el fin de comprender el impacto que tienen las emociones en el estudiante, se analizaron varias teorías. En ellas se expone la existencia de procesos superiores que tienen control sobre los diferentes estados y experiencias de la mente: las emociones. Las emociones, del mismo modo que los procesos cognitivos, son sistemas adaptativos capaces de orquestar y controlar cada respuesta cognitiva (de pensamiento o respuesta corporal). Cada emoción es capaz de ajustar sus parámetros, de modo que cuando se detecta una situación, se envía una señal a los diferentes subprocesos y respuestas cognitivas a través de reacciones químicas y eléctricas. De esta manera, se controlan, procesos como los de percepción, atención, aprendizaje, memoria, reacciones fisiológicas (como frecuencia cardiaca, función endocrina, función inmunológica), e incluso variables reguladoras (como puede ser el autoestima). Los argumentos detallados frente a estas posiciones pueden encontrarse en [2, 8, 10–12].

Nuestro interés en comprender las teorías, de psicología y neurociencia, ligadas a un comportamiento psicológico y a las interacciones afectivo-cognitivas, tiene como propósito entender las implicaciones de las emociones en las problemáticas adaptativas implícitas en los procesos de aprendizaje. Esta visión comprende la importancia de la educación como medio formativo de adaptaciones en los procesos de control cognitivo, motivacional y de experiencia social y emocional.

Los estudios en inteligencia emocional (IE) han demostrado que los estudiantes con más coeficiente emocional [5] manejan menores niveles de estrés, menor ansiedad social y depresión; adicionalmente, cuando se enfrentan a tareas complejas, perciben los estresores como menos amenazantes. Incluso otros estudios evidencian una fuerte correlación entre la inteligencia emocional y el rendimiento académico del estudiante.

---

<sup>4</sup> <http://www.affective-sciences.org/researchmaterial>

Desde la perspectiva psicológica, una de las teorías más conocidas considera que las emociones corresponden a diferentes reacciones fisiológicas instantáneas y discretas que pueden dividirse en familias de emociones básicas [4]. Por esto define seis emociones básicas en el ser humano: sorpresa, asco, tristeza, ira, miedo, y alegría. En el ámbito educativo, sin embargo, las emociones consideradas relevantes son nueve: alegría, esperanza, orgullo, gratitud, disfrute, ira, frustración, ansiedad y vergüenza [9]. Estas emociones fueron las seleccionadas para el diseño de nuestro conjunto de datos.

De acuerdo a esto lo anterior es fundamental que los estudiantes reconozcan lo que están sintiendo, lo asimilen, lo comprendan y lo regulen. Este proyecto busca apoyar el primer momento: el reconocimiento de la emoción.

## 2.2. Trabajos relacionados

Actualmente, el interés de las investigaciones enfocadas en la clasificación de emociones ha crecido constantemente.

Uno de los aspectos más importantes para generar un sistema autónomo capaz de clasificar la emoción expresada textualmente bajo los escenarios del aprendizaje automático, son las fuentes de datos usadas para entrenar las máquinas. En [13] se describen exhaustivamente los más diversos y esenciales conjuntos de datos en el área bajo diferentes contextos (análisis de Twitter, opiniones de libros y películas, Blogs, entre otros).

En el ámbito educativo, el conjunto de datos *ISEAR* expone la idea de describir una situación textualmente donde se experimentó cierta emoción. Fue construido por un grupo de psicólogos, que les preguntaron a un grupo de estudiantes en qué situaciones habían experimentado cada una de las siete emociones consideradas en el estudio: alegría, miedo, ira, tristeza, disgusto, vergüenza y culpa). La encuesta terminó reuniendo cerca de 3000 respuestas de participantes en 37 países.

En el contexto general, las investigaciones y aportes actuales de fuentes de datos para la clasificación de emociones textualmente, resaltan el uso de *SemEval-2007 task #14*, debido a que es un estándar para la prueba de modelos de clasificación emocionales.

Los trabajos análogos relevantes se consolidaron en la tabla 1. Los trabajos [3] y [7] exhibieron modelos con el mejor desempeño de validación respecto a SemEval y ISEAR. El mejor modelo entrenado y validado con ISEAR, fue una máquina de soporte vectorial, con un desempeño F1 de 0,68.

## 3. Metodología

La metodología definida en el trabajo pretende permitir validar varias presunciones sobre el potencial de los modelos para la detección de emociones:

1. Construir un clasificador para texto en inglés. con el objetivo de apropiarnos de las estrategias aplicadas en trabajos relacionados exitosos y familiarizarnos con las bases de de datos más usadas en clasificación de emociones

Methodo - Algoritmo		DataSet		Resultados		Personaje	
		Entrenamiento	Validación	Accuracy	F1		
ML - Supervisado	Redes Bayesianas (NB)	LiveJournal Blog	ISEAR	-	<b>0,13</b>	C. Strapparava and R. Mihalcea [2008]	
Basado en Conocimiento	WN - Precencia de Afecto (Lexicon)	-		-	<b>0,04</b>		
Basado en Conocimiento	Análisis Semántico Latente (LSA)	-		-	<b>0,18</b>		
Basado en Conocimiento	Basado en Reglas (RB)	-		SemEval 2007	0,89	<b>0,09</b>	F. R. Chaumartin [2007]
ML - Supervisado	Redes Bayesianas (NB)	-			0,67	<b>0,67</b>	T. Danisman and A. Alpkocak [2008]
ML - Supervisado	Maquina de Soporte Vectorial (SVM)	ISEAR			0,70	<b>0,68</b>	
ML - Supervisado	Basado en Conocimiento (KB - VSM)	-			0,36	<b>0,29</b>	
Basado en Conocimiento	Modelo de espacio vectorial (VSM)	-		-	<b>0,51</b>	S. Kim et al. [2010]	
ML - Supervisado	Modelo Oculuto de Markov (HMM)	ISEAR		ISEAR	-	<b>0,35</b>	D. T. Ho and T. H. Cao [2012]
Basado en Conocimiento	Basado en Conocimiento (KB - VSM)	-			-	<b>0,37</b>	S. Kim et al. [2010]

Cuadro 1: Rendimiento de los mejores modelos de clasificación de emociones

(ISEAR y SemEval), de manera que se logre componer un clasificador con un desempeño comparable al de dichos trabajos.

2. Construir un clasificador para texto en español. El propósito es explorar las implicaciones de cambio de lenguaje, la meta es generar un conjunto de datos de clasificación de emociones en español (ISEAR traducido) y un clasificador con desempeño comparable al de la etapa anterior.
3. Construir un clasificador para texto de la comunidad del Programa de Ingeniería de Sistemas de la Escuela Colombiana de Ingeniería Julio Garavito, con el propósito de tener el primer conjunto de datos y el primer clasificador de emociones para esta comunidad.

Para construir el clasificador, se propone el entrenamiento de cuatro máquinas distintas (máquina de soporte vectorial (SVM), red de Bayes (NB), red neuronal (NN), lexicón). El proceso de entrenamiento (con un 80 % de los datos) se hace a partir de ISEAR y ISEAR traducido, y se establece su respectiva validación cruzada (con el 20 % restante). Luego, se procede a validar la hipótesis de alta dependencia del contexto, a partir de las pruebas con los datos de SemEval. Adicionalmente, se propone la construcción de una máquina compuesta por los cuatro clasificadores anteriormente mencionados.

## 4. Resultados obtenidos

### 4.1. Exploración y pre-procesamiento de datos

De ISEAR, se tomaron 2 de las **43** características, *SIT* y *EMOT*; *SIT* proporciona información de la frase contestada por el estudiante, *EMOT* contiene la etiqueta a la que la frase responde. Con el fin de ver la distribución de las emociones y encontrar algún sesgo entre ellas, se exploró el conjunto de datos y se descubrió que el número de ejemplares es igual para cada una de las emociones (1095 para un total de 7665). Por otro lado, SemEval contiene 2 conjuntos de datos separados, *trial* y *test*, con 250 y 1000 datos respectivamente. En la Figura 1 presenta una síntesis de los ejemplares de ISEAR y SemEval.

Durante el pre-procesamiento se realizaron varias etapas con el fin de resolver diferentes problemáticas.

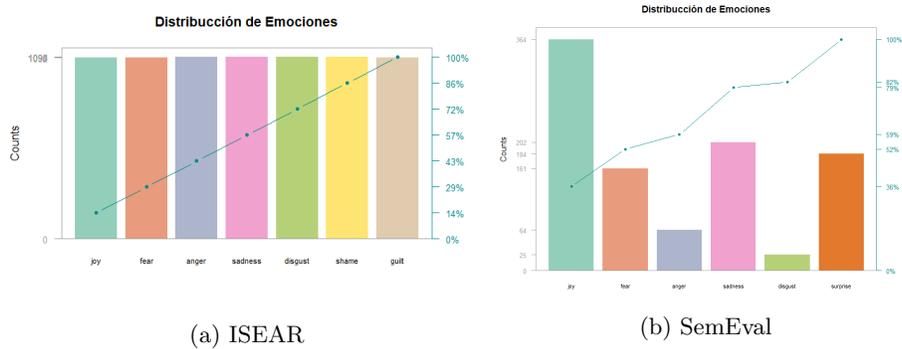


Figura 1: Distribución Inicial de Datos

1. Al haber en el texto palabras sin sentido o palabras vacías (e.g "the", ".and", "was", ...) se eliminaron del conjunto de datos.
2. Se eliminaron caracteres especiales y símbolos de puntuación y acentuación; adicionalmente, se suprimieron los números, los cuales para el propósito de este trabajo no son necesarios.
3. Aplicamos el proceso de llevar las palabras a su raíz (*stemming*).

Luego de realizar el pre-procesamiento, se compararon algunas de las particularidades de los conjuntos de datos con el preprocesado. Las comparaciones en el mapa de palabras (Figura 2), permite distinguir de una mejor manera las palabras emotivas sin el ruido que ocasionaban las palabras vacías.

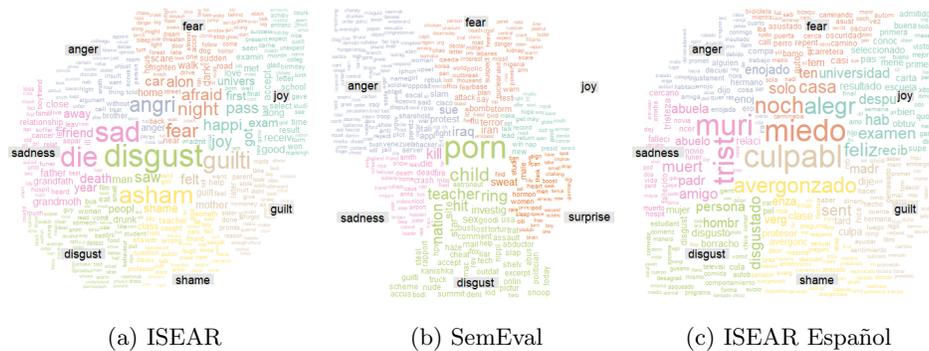


Figura 2: Mapa de Palabras después del pre-procesamiento

## 4.2. Modelos simples

Se realizó el entrenamiento de 3 modelos de aprendizaje supervisado, con el fin de evaluar qué modelo nos proporciona la mejor precisión. Adicionalmente al modelo basado en conocimiento externo, el lexicón NRC <sup>5</sup> se le realizó la evaluación.

Con el fin de realizar los entrenamientos de los modelos, se partió el conjunto de datos en 80 % para entrenamiento y 20 % para test con ISEAR. Para la SVM con un kernel lineal, el mejor learning rate fue 0,5. Luego de realizar el pre-procesamiento, el número de datos con los que se entrenó fueron 6128 con 12290 características. En cuanto a la NN, su arquitectura fueron 2 capas ocultas con 20 y 10 neuronas respectivamente, y por ultimo se entrenó la NB con el mismo número de datos.

Con el fin de realizar pruebas a nuestros modelos se realizaron 2 tipos de análisis. El primero consistió en revisar cómo se comportaba al entrenar y probar con ISEAR. La segunda aproximación se enfocó en realizar el entrenamiento con ISEAR y probar con SemEval arrojando una baja precisión.

Se realizaron pruebas de validación cruzada con ISEAR (Tabla 2), para lo cual obtuvimos que el modelo que mejor se comportó fue la máquina de soporte vectorial que obtuvo casi un 53 % de precisión y un F1 de 0,54.

Metodo		DataSet		PreProcesamiento	Sparse	Dim	Representación de Características	Resultados			
Algoritmo	Parametros	Entrenamiento	Validación					Accuracy	Precision	Recall	F1
Maquina de Soporte Vectorial (SVM)	Kernel Lineal $\alpha : 0.05$	ISEAR	ISEAR	Palabras parada / Stemming	0,999	7661 x 12290	Bag of Words [Frecuencias]	0,5339	0,5777	0,5339	<b>0,5414</b>
Redes Neuronales (NN)	Capa Oculta : 20 - 10			Palabras parada / Stemming	0,999	7661 x 12290	Bag of Words [Frecuencias]	0,4211	0,4317	0,4213	<b>0,4216</b>
Redes Bayesianas (NB)	-			Palabras parada / Stemming	0,99	7361 x 376	Bag of Words [Frecuencias]	0,33	-	-	-
Basado en Conocimiento (KB - NRC)	Emociones Comunes NRC			Palabras parada / Stemming	0,999	7661 x 12290	Tokenization	0,4142	0,4187	0,4164	<b>0,3935</b>
Máxima entropía	Emociones Comunes NRC			Palabras parada / Stemming	-	7661 x 12290	DocumentTermMatrix	0,1486	0,1652	0,1601	<b>0,1605</b>

Cuadro 2: Resultados entrenamiento y validación con ISEAR.

Según los análisis ROC que se realizaron, las máquinas se comportan mejor realizando predicciones de datos relacionados con las emociones de disfrute y miedo.

Para el modelo SVM (Figura 3a), con un Learning Rate de 0,5, como se mencionó, fue el que mejor se comportó y se observó que las predicciones realizadas por disfrute y miedo fueron las más precisas.

De la misma manera que en la SVM, en las redes neuronales (Figura 3b), con 100 neuronas en capa oculta, se observa un comportamiento similar para las emociones de disfrute y miedo, lo que nos lleva a pensar que los patrones de expresión para estas emociones son más claros.

Durante la evaluación con Redes Bayesianas (Figura 3c), el modelo no se comportó adecuadamente, incluso luego de normalizar los datos, el accuracy y

<sup>5</sup> <https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

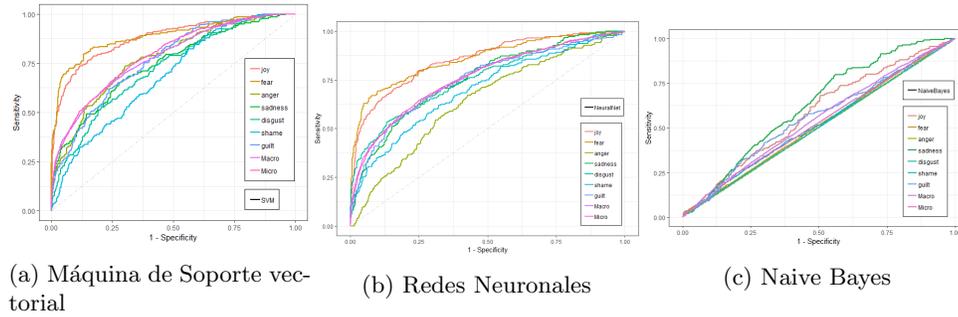


Figura 3: Análisis ROCR por cada modelo entrenado.

F1 del modelo no subió tanto como se esperaba al inicio. Y aunque la emoción que mejor se comportó fue la tristeza, los resultados de la red no fueron satisfactorios.

Hicimos pruebas con SemEval teniendo a ISEAR como conjunto de datos de entrenamiento. Luego de realizar las pruebas correspondientes por cada modelo, notamos un muy bajo rendimiento (Tabla 3). Lo anterior se explica porque los conjuntos de datos son de un dominio distinto; por lo tanto, al entrenar con los datos de ISEAR, no se abstrae la suficiente información para tratar los datos de SemEval.

Metodo		DataSet		PreProcesamiento	Sparse	Dim	Representación de Características	Resultados			
Algoritmo	Parametros	Entrenamiento	Validación					Accuracy	Precision	Recall	F1
Máquina de Soporte Vectorial (SVM)	Kernel Lineal $\alpha : 0.05$	ISEAR	SEM Eval	Palabras parada / Stemming	0,999	809 x 2159	Bag of Words [Frecuencias]	0,2484	0,2480	0,2217	<b>0,2341</b>
Redes Neuronales (NN)	Capa Oculta : 20 - 10			Palabras parada / Stemming	0,999	809 x 2159	Bag of Words [Frecuencias]	0,1569	0,1702	0,1855	<b>0,1775</b>
Redes Bayesianas (NB)	-			Palabras parada / Stemming	0,99	809 x 2159	Bag of Words [Frecuencias]	0,0700	-	0,2000	-
Basado en Conocimiento (KB - NRC)	Emociones Comunes NRC			Palabras parada / Stemming	0,99			-	-	-	<b>0,0550</b>

Cuadro 3: Resultados de pruebas con SemEval.

El modelo con mejores resultados continúa siendo la de Soporte Vectorial, con un accuracy de 24.84 %.

### 4.3. Modelos compuestos

Con el fin de mejorar los resultados, exploramos tres aproximaciones, aún en revisión, de modelos compuestos por los descritos anteriormente.

- **Máquina por votación:** el resultado es la emoción seleccionada por el mayor número de máquinas.
- **Máquina por probabilidades:** el resultado es la emoción que tiene mayor probabilidad de las seleccionadas por todas las máquinas.

- **Máquina por expertos:** el resultado se calcula considerando el desempeño de las diferentes máquinas durante el entrenamiento, considerando su F1 y ROCR.

Estos modelos mejoraron un 10 % (Figura 4) la precisión respecto a la clasificación de los modelos individuales en SemEval.

Metodo	DataSet		PreProcesamiento	Resultados			
	Entrenamiento	Validación		Accuracy	Precision	Recall	F1
Composed Simple Machine	ISEAR	SEM Eval	Palabras parada / Stemming	0,2140	0,2090	0,1771	<b>0,1917</b>
Composed Probabilities Machine			Palabras parada / Stemming	0,0840	0,1755	0,1522	<b>0,1630</b>
Composed Expert Machine			Palabras parada / Stemming	-	-	-	<b>0,3400</b>

Cuadro 4: Resultados entrenamiento-evaluación máquinas compuestas ISEAR-SemEval.

#### 4.4. ISEAR traducido

Adicionalmente, propusimos la hipótesis de que la traducción del conjunto de datos no afectaría el rendimiento de los modelos diseñados. Para probar esta hipótesis, traducimos ISEAR y entrenamos el modelo con mejor resultado en inglés: la máquina de soporte vectorial. El desempeño de este nuevo modelo fue equivalente al logrado con ISEAR original.

#### 4.5. Academic Survey on Emotional Reactions in Computational Learning Processes (ASER-CLP)

Durante la investigación generamos un conjunto de datos en español para conocer el estado de inteligencia emocional de estudiantes de pregrado de la Escuela Colombiana de Ingeniería Julio Garavito. Para esto diseñamos una encuesta de 9 preguntas con 3 tres enfoques diferentes: (i) identificar el perfil emocional, (ii) identificar descripciones de situaciones (pares emoción - texto) con mayor recordación emocional positivas y negativas, y (iii) identificar expresión de emociones (pares emoción - texto) ante situaciones críticas comunes en procesos de enseñanza.

La encuesta la aplicamos a 300 estudiantes del Programa de Ingeniería de Sistemas y obtuvimos el 63 % de respuestas. Los resultados fueron los siguientes: (i) la frecuencia de las emociones en los estudiantes (Figura 4) (ii) 510 ejemplares de emoción - texto de las situaciones de mayor recordación (iii) 510 ejemplares de emoción - texto de reacciones a situaciones comunes.

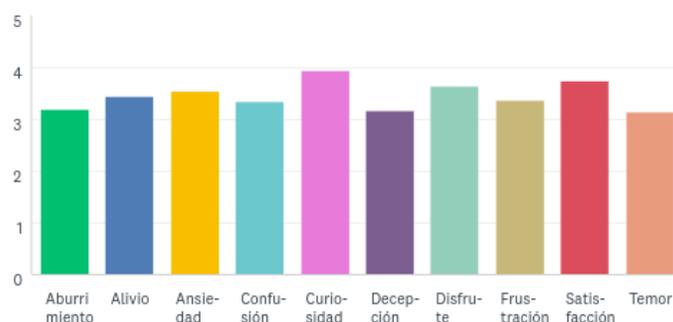


Figura 4: Frecuencia de las emociones

Las emociones que más expresaron los estudiantes en (ii) y (iii) (satisfacción, frustración y alivio) son las más experimentadas en procesos educativos de acuerdo con estudios en el área; adicionalmente, en la muestra, a diferencia de otros estudios, la ansiedad es también de las más emociones más expresadas.

## 5. Conclusiones y trabajo futuro

Las máquinas entrenadas durante la investigación se aproximaron bastante a lo esperado considerando los resultados de trabajos análogos; sin embargo, con el fin de lograr mejor desempeño, hay varias alternativas por explorar. Entre ellas, para el pre-procesamiento y representación de características, se deben analizar otras aproximaciones para representar de los datos (como TF-IDF) y usar estrategias para tratar sinónimos, antónimos y negaciones con el fin de potenciar el contenido semántico de los ejemplares. Adicionalmente, se recomienda que al desarrollar trabajos de minería de texto, se considere explícitamente el contexto del problema, incluso abreviaciones y modismos.

En un próximo trabajo asociado a este proyecto, se busca perfeccionar los modelos de clasificación para contexto educativo en español contemplando ISEAR traducido y ASER-CLP (el conjunto inicial de datos de la institución). De igual modo, se piensa hacer el despliegue del modelo entrenado mediante un micro-servicio que dado un texto de contexto educativo clasifique la emoción para que las instituciones, educadores y alumnos puedan conocer el estado emocional de estos últimos y lo tengan en cuenta en el diseño de los procesos de aprendizaje.

## Referencias

1. Caridakis, G., Castellano, G., Kessous, L., Raouzaïou, A., Malatesta, L., Asteriadis, S., Karpouzis, K.: Multimodal emotion recognition from expressive faces, body gestures and speech. In: Boukis, C., Pnevmatikakis, A., Polymenakos, L. (eds.) *Artificial Intelligence and Innovations 2007: from Theory to Applications*. pp. 375–388. Springer US, Boston, MA (2007)

2. da Cosmides, I., Tooby, J.: Dissecting the computational architecture of social inference mechanisms. *Ciba Foundation symposium* (1997)
3. Danisman, T., Alpkocak, A.: Feeler: Emotion classification of text using vector space model. In: *AISB 2008 Convention Communication, Interaction and Social Intelligence*. vol. 1, p. 53 (2008)
4. Ekman, P.: An argument for basic emotions. *Cognition and Emotion* **6**(3-4), 169–200 (1992). <https://doi.org/10.1080/02699939208411068>, <https://doi.org/10.1080/02699939208411068>
5. Extremera Pacheco, N., Fernández-Berrocal, P.: The role of students' emotional intelligence: Empirical evidence. *SicElo* **6**(2) (2004)
6. Hancock, J.T., Landrigan, C., Silver, C.: Expressing emotion in text-based communication. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 929–932. CHI '07, ACM, New York, NY, USA (2007). <https://doi.org/10.1145/1240624.1240764>, <http://doi.acm.org/10.1145/1240624.1240764>
7. Kim, S.M., Valitutti, A., Calvo, R.A.: Evaluation of unsupervised emotion models to textual affect recognition. In: *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*. pp. 62–70. Association for Computational Linguistics (2010)
8. Panksepp, J.: *Affective neuroscience: The foundations of human and animal emotions*. Oxford university press (2004)
9. Phye, G.D., Schutz, P., Pekrun, R.: *Emotion in education*. Elsevier (2011)
10. Tooby, J.: The emergence of evolutionary psychology. *Emerging syntheses in science* pp. 67–76 (1985)
11. Tooby, J., Cosmides, L.: On the universality of human nature and the uniqueness of the individual: The role of genetics and adaptation. *Journal of personality* **58**(1), 17–67 (1990)
12. Tooby, J., Cosmides, L.: The past explains the present: Emotional adaptations and the structure of ancestral environments. *Ethology and sociobiology* **11**(4-5), 375–424 (1990)
13. Yadollahi, A., Shahraki, A.G., Zaiane, O.R.: Current state of text sentiment analysis from opinion to emotion mining. *ACM Computing Surveys (CSUR)* **50**(2), 25 (2017)