



UAI
UNIVERSIDAD ADOLFO IBÁÑEZ



Caso Aplicado NLP

**Detección temprana del riesgo de depresión
mediante Procesamiento de Lenguaje Natural**

Enero 2026

Integrantes del grupo:

Bastián Velasco

René Ruiz

Pablo Gutiérrez

Daniel Cabedo



Pregunta 01: ¿Qué “consideraciones” (cuidados) habría que tener para el desarrollo e implementación de la solución?

Consideración 01: Propuesta de Valor y Uso Clínico Responsable

La primera consideración clave para el desarrollo e implementación de la solución es definir con claridad su **propuesta de valor**, estableciendo explícitamente que el sistema de NLP tiene un rol de **apoyo clínico** y no de reemplazo del diagnóstico profesional. La solución debe entregar como resultado un **score continuo de riesgo de depresión**, que permita clasificar a los pacientes en distintos **niveles de riesgo** (por ejemplo, bajo, medio y alto), habilitando acciones diferenciadas como priorización de atención, activación de protocolos preventivos o derivación oportuna. Este enfoque busca aumentar la capacidad de detección temprana sin introducir automatismos que puedan comprometer la responsabilidad clínica. La integración con los flujos existentes del sistema de salud —como las fichas clínicas electrónicas y los procesos de triaje— es esencial para asegurar adopción y utilidad real en la práctica.

Supuestos que sustentan esta consideración:

Se asume que los profesionales de la salud mental participan activamente en el diseño y validación del sistema, y que existe consenso institucional en que el modelo entrega recomendaciones basadas en riesgo, pero no diagnósticos definitivos. Se asume además que el score de riesgo será acompañado de lineamientos clínicos claros sobre las acciones a tomar en cada nivel, evitando interpretaciones erróneas. Finalmente, se presupone que los usuarios finales recibirán capacitación adecuada para comprender las capacidades y limitaciones del sistema, fortaleciendo la confianza y el uso responsable de la herramienta.



Pregunta 01: ¿Qué “consideraciones” (cuidados) habría que tener para el desarrollo e implementación de la solución?

Consideración 02: Datos Clínicos, Privacidad y Completitud de Diagnósticos

La segunda consideración crítica se relaciona con la **gestión de los datos**, dado que la solución se construye sobre información clínica altamente sensible. Las transcripciones de sesiones paciente–profesional contienen contenido íntimo y contextual que exige un manejo estricto de privacidad, seguridad y ética. Es indispensable contar con **consentimiento informado explícito** para la grabación y transcripción de las sesiones, así como con protocolos robustos de anonimización y control de acceso. Adicionalmente, dado que solo un subconjunto de pacientes cuenta con diagnóstico de depresión, se vuelve necesario diseñar un **proceso paralelo de completitud diagnóstica**, en el cual los profesionales de la salud puedan revisar y etiquetar retrospectivamente casos sin diagnóstico en la base histórica, fortaleciendo la calidad del set supervisado para entrenamiento y evaluación del modelo.

Supuestos que sustentan esta consideración:

Se asume el cumplimiento estricto de la normativa vigente en protección de datos de salud y derechos del paciente, incluyendo estándares de confidencialidad y trazabilidad. Se asume también la disponibilidad y disposición de profesionales clínicos para participar en procesos de etiquetado adicional, bajo esquemas de carga de trabajo acotados y priorizados. Finalmente, se presupone que la institución cuenta con capacidades técnicas y legales para implementar mecanismos de anonimización que reduzcan riesgos de reidentificación sin comprometer el valor clínico del texto.



Pregunta 01: ¿Qué “consideraciones” (cuidados) habría que tener para el desarrollo e implementación de la solución?

Consideración 03: Riesgo Clínico, Explicabilidad y Sostenibilidad del Sistema



La tercera consideración fundamental es abordar los **riesgos clínicos y operacionales** asociados al uso de modelos de clasificación de texto en un contexto de salud mental. Errores como **falsos negativos** (no detectar un caso de alto riesgo) o **falsos positivos** (sobre-reaccionar ante un riesgo bajo) tienen impactos asimétricos y potencialmente graves, por lo que el diseño del sistema debe priorizar criterios clínicos por sobre métricas puramente técnicas. En este contexto, la **explicabilidad del modelo** es clave para que los profesionales comprendan por qué un paciente obtiene determinado score de riesgo, facilitando la validación humana y la confianza en la herramienta. Asimismo, la solución debe ser sostenible en el tiempo, considerando escalabilidad, costos de mantenimiento, reentrenamiento periódico y adaptación a cambios en el lenguaje clínico y social.



Supuestos que sustentan esta consideración:

Se asume que la institución prioriza la seguridad del paciente por sobre la optimización extrema del desempeño predictivo, aceptando trade-offs en precisión si ello mejora la interpretabilidad y la robustez del sistema. Se asume también que existen recursos mínimos para mantener la infraestructura tecnológica y realizar monitoreo continuo del desempeño del modelo en producción. Finalmente, se presupone un marco regulatorio que exige transparencia y auditabilidad en sistemas de alto riesgo, alineando el diseño del modelo con principios de explicabilidad y responsabilidad.





Pregunta 02: ¿Qué podríamos hacer para “aumentar” la probabilidad de una implementación exitosa?

Consideración 01: Desarrollo Iterativo y Aprendizaje Recursivo del Modelo

Para aumentar la probabilidad de una implementación exitosa, es clave adoptar un enfoque de **desarrollo iterativo con aprendizaje recursivo**, donde el modelo evolucione progresivamente a partir del uso real. La solución debe combinar **aprendizaje supervisado** con estrategias que permitan incorporar nuevos datos de manera continua. En particular, las transcripciones que ya cuentan con diagnóstico confirmado por profesionales deben alimentar ciclos sucesivos de reentrenamiento, mientras que los casos con diagnósticos faltantes pueden priorizarse mediante **active learning**, enfocando el esfuerzo clínico en aquellos ejemplos donde el modelo presenta mayor incertidumbre. Este enfoque permite mejorar el desempeño del sistema en contextos de datos limitados y asegura que el modelo se adapte a cambios en la población y en las prácticas clínicas.

Supuestos que sustentan esta consideración:

Se asume la existencia de una infraestructura mínima para versionar modelos, datos y experimentos, garantizando trazabilidad y reproducibilidad. Se asume también la disponibilidad periódica de profesionales para validar y etiquetar nuevos casos priorizados por el sistema. Finalmente, se presupone que la organización acepta un despliegue gradual, donde el modelo mejora progresivamente y no se espera un desempeño óptimo desde la primera versión productiva.



Pregunta 02: ¿Qué podríamos hacer para “aumentar” la probabilidad de una implementación exitosa?

Consideración 02: Gobernanza de Datos y Evaluación Clínica Integral

Una segunda acción clave para aumentar la probabilidad de éxito es establecer una **gobernanza de datos sólida** acompañada de un esquema de **evaluación que trascienda las métricas técnicas tradicionales**. Dado el carácter sensible y crítico del problema, el desempeño del modelo debe evaluarse no solo en términos de accuracy o F1-score, sino considerando recall para casos de alto riesgo, estabilidad temporal del modelo, sesgos demográficos y utilidad clínica percibida. La gobernanza debe definir con claridad quién puede acceder a los datos, con qué fines, bajo qué controles y cómo se auditan los usos del sistema. Esto es especialmente relevante en un entorno donde el modelo influye en decisiones de priorización clínica.

Supuestos que sustentan esta consideración:

Se asume que la institución cuenta con capacidades organizacionales para definir roles, responsabilidades y procesos de auditoría de datos y modelos. Se asume también que existen métricas clínicas consensuadas que permiten evaluar el impacto real del sistema en la práctica (por ejemplo, reducción de tiempos de detección o mejora en cobertura preventiva). Finalmente, se presupone acceso a información demográfica suficiente para monitorear y mitigar sesgos sin vulnerar la privacidad de los pacientes.



Pregunta 02: ¿Qué podríamos hacer para “aumentar” la probabilidad de una implementación exitosa?

Consideración 03: Implementación Gradual y Gestión del Cambio Organizacional

La tercera acción para aumentar la probabilidad de una implementación exitosa es abordar explícitamente la **gestión del cambio organizacional**, reconociendo que la adopción de soluciones de NLP en salud mental es un proceso social y cultural, además de técnico. Se recomienda un **despliegue gradual**, comenzando con pilotos controlados en centros específicos, donde se pueda ajustar el uso del score de riesgo, los umbrales de alerta y los protocolos asociados antes de una expansión a mayor escala. La capacitación de los profesionales debe enfocarse en clarificar qué puede y qué no puede hacer el sistema, cómo interpretar los niveles de riesgo y cómo integrar la herramienta en la toma de decisiones clínicas diarias. La retroalimentación continua de los usuarios finales es clave para ajustar tanto el modelo como los procesos que lo rodean.

Supuestos que sustentan esta consideración:

Se asume el apoyo institucional del Ministerio de Salud para promover la adopción progresiva y respaldar a los equipos clínicos durante el proceso de cambio. Se asume también la identificación de líderes clínicos que actúen como referentes y facilitadores del uso de la herramienta. Finalmente, se presupone que la organización está dispuesta a invertir tiempo en capacitación y acompañamiento, entendiendo que la construcción de confianza en sistemas de apoyo basados en IA es un proceso gradual y acumulativo.



Conclusiones

El desarrollo de una solución de NLP para la detección temprana del riesgo de depresión en el sistema público de salud representa una oportunidad significativa para mejorar la prevención y priorización de la atención en salud mental, pero también implica desafíos técnicos, éticos y organizacionales de alta complejidad. A lo largo de la propuesta se ha enfatizado que el valor de la solución no reside únicamente en el desempeño del modelo, sino en su correcta integración como herramienta de apoyo clínico, basada en un score de riesgo interpretable y accionable, siempre bajo supervisión profesional.

Asimismo, se destaca la centralidad de los datos: el consentimiento informado, la protección de la privacidad y la completitud diagnóstica son condiciones habilitantes para cualquier implementación responsable. El enfoque de aprendizaje recursivo, apoyado en la participación activa de los profesionales de la salud, permite enfrentar contextos de datos limitados y asegurar la mejora continua del sistema. Finalmente, la probabilidad de éxito depende en gran medida de una implementación gradual, con gobernanza clara, evaluación clínica integral y una gestión del cambio que promueva confianza y adopción sostenida. En conjunto, estas consideraciones alinean la propuesta con los principios de un NLP responsable, clínicamente relevante y sostenible en el tiempo.