

ÍNDICE DE CONTENIDOS

| | |
|---|----|
| RESUMEN..... | 11 |
| ABSTRACT..... | 13 |
| 1 MARCO TEÓRICO | 15 |
| 1.1 SISTEMAS MONOAMINÉRGICOS Y RECEPTORES NICOTÍNICOS | 17 |
| 1.1.1 Proteínas del sistema Monoaminérgico | 17 |
| 1.1.2 Receptores Nicotínicos de Acetilcolina | 18 |
| 1.2 SISTEMAS MONOAMINÉRGICOS Y RECEPTORES NICOTÍNICOS, SU IMPORTANCIA EN TRASTORNOS NEURODEGENERATIVOS | 20 |
| 1.3 TÓPICOS FARMACOLÓGICOS RELEVANTES..... | 22 |
| 1.3.1 Promiscuidad de Fármacos..... | 22 |
| 1.3.2 Poliselectividad de Fármacos..... | 23 |
| 1.4 ALGORITMOS DE COMPARACIÓN DE SITIOS DE UNIÓN | 23 |
| 1.5 MINERÍA DE DATOS | 25 |
| 1.5.1 Aprendizaje Automático (“ <i>Machine Learning</i> ”) | 25 |
| 1.5.2 Estudios de farmacología con aprendizaje automático | 26 |
| 2 PREGUNTA TÉCNICA..... | 30 |
| 3 OBJETIVOS..... | 32 |
| 3.1 OBJETIVO GENERAL:..... | 32 |
| 3.2 OBJETIVOS ESPECÍFICOS: | 32 |
| 4 METODOLOGÍA..... | 33 |
| 4.1 GENERAR UN SET DE REFERENCIA..... | 34 |
| 4.1.1 Construcción del set de referencia..... | 34 |
| 4.1.2 Extracción de sitios de unión..... | 35 |

| | |
|---|----|
| 4.2 CARACTERIZAR LOS COMPLEJOS | 36 |
| 4.2.1 Atributos Geométrico..... | 37 |
| 4.2.2 Hidrofobicidad | 42 |
| 4.2.3 Composición Aminoacídica | 42 |
| 4.2.4 Composición Atómica por Capa..... | 43 |
| 4.2.5 Energía..... | 44 |
| 4.2.6 Radios de Sitio | 45 |
| 4.3 ANÁLISIS ESTADÍSTICO DE ATRIBUTOS | 45 |
| 4.4 ENTRENAMIENTO Y EVALUACIÓN DE MODELOS | 46 |
| 4.4.1 Algoritmos de clasificación | 46 |
| 4.4.2 Medidas de evaluación de performance..... | 50 |
| 5 RESULTADOS | 53 |
| 5.1 SET DE REFERENCIA | 53 |
| 5.2 CARACTERIZAR LOS COMPLEJOS | 53 |
| 5.3 ANÁLISIS ESTADÍSTICO | 55 |
| 3.1 Normalización de los atributos | 56 |
| 5.4 EVALUACIÓN DE EFICIENCIA DE DISTINTOS ALGORITMOS..... | 57 |
| 5.4.1 Resultados Iniciales | 57 |
| 5.4.2 RESULTADOS CON SET DE REFERENCIA BALANCEADO..... | 60 |
| 5.4.3 DIFERENCIAS ENTRE SET CON Y SIN BALANCE | 65 |
| 5.4.4 OPERATIVIDAD DEL MODELO | 69 |
| 6 DISCUSIÓN..... | 74 |
| 7 CONCLUSIONES | 77 |
| 8 REFERENCIAS | 78 |
| 9 ANEXOS..... | 83 |

| | |
|--|-----|
| Anexo 1.a: Script para generar una matriz con los aminoácidos por distancia al centro de masa del fármaco. | 83 |
| Anexo 1.b: Script para ordenar los aminoácidos en relación a la distancia de los mismos al centro de masa del fármaco. | 83 |
| Anexo 2: Script para recopilar los datos relevantes de los cuatro aminoácidos más cercanos al centro de masa del fármaco. | 84 |
| Anexo 3: Script para obtener distancias y ángulos. | 85 |
| Anexo 4.a: Script para generar el archivo. | 88 |
| Anexo 4.b: Script para obtener el volumen con el módulo msms. | 88 |
| Anexo 5.a: Script para calcular torsión entre los cuatro aminoácidos más cercanos. | 89 |
| Anexo 5.b: Script para calcular ángulos ϕ y ψ | 89 |
| Anexo 6: Escalas de hidrofobicidad. | 91 |
| Anexo 7: Script para calcular hidrofobicidad y composición aminoacídica. | 92 |
| Anexo 8: Script para calcular átomos por capas | 96 |
| Anexo 9: Script para ocupar <i>Foldx</i> | 99 |
| Anexo 10: Escala de energía de captación por aminoácido. | 100 |
| Anexo 11: Script para calcular energía de captación | 100 |
| Anexo 12: Script para calcular radios de sitios | 102 |
| Anexo 13: Script para normalizar los atributos | 103 |
| Anexo 14: Indicadores estadísticos para atributos geométricos. | 105 |
| Anexo 14.a: Distancias. | 105 |
| Anexo 14.b: Ángulos | 106 |
| Anexo 14.c: Volumen | 106 |
| Anexo 14.d: Torsiones | 107 |
| Anexo 15: Indicadores estadísticos para la composición aminoacídica. | 107 |

| | |
|---|-----|
| Anexo 16: Indicadores estadísticos para la composición atómica..... | 108 |
| Anexo 17: Indicadores estadísticos para los radios del sitio | 108 |

ÍNDICE DE FIGURAS

| | |
|---|----|
| Figura 1: Proceso de desarrollo de un fármaco..... | 15 |
| Figura 2 a. Estructura química del NE b. Estructura química del DA c. Estructura química del 5-HT | 18 |
| Figura 3: Estructura tridimensional del dominio de unión de ligando de nAChRs..... | 19 |
| Figura 4: Proceso de selección de fármacos candidatos para 5HT ₃ | 29 |
| Figura 5: modelo de clasificación de aprendizaje automático..... | 33 |
| Figura 6: Composición del set de entrenamiento..... | 35 |
| Figura 7: Sitio de unión analizado..... | 36 |
| Figura 8: Distancias medidas para los 4 aminoácidos más cercanos al centro de masa del fármaco. | 38 |
| Figura 9: Ángulos medidos para los 4 aminoácidos más cercanos al centro de masa del fármaco. | 39 |
| Figura 10: Representación de torsión o ángulo diedro. | 40 |
| Figura 11: Representación de la forma de operar de SVM..... | 48 |
| Figura 12: Proceso de clasificación por <i>KNN</i> | 49 |
| Figura 13: Ejemplo de matriz de confusión..... | 50 |
| Figura 14: Curvas ROC para el set reducido con diversos valores de <i>smote</i> | 66 |
| Figura 15: Curvas ROC del set completo con diferentes valores de <i>smote</i> | 68 |
| Figura 16: Propuesta de operatividad para el modelo creado con el set reducido. ... | 70 |
| Figura 17: Propuesta de operación para el modelo generado con el set completo. ... | 72 |

ÍNDICE DE TABLAS

| | |
|--|-----|
| Tabla 1: Distribución de ejemplos set reducido. | 53 |
| Tabla 2: Distribución de ejemplos set completo..... | 53 |
| Tabla 3: Descripción de las modificaciones de los atributos..... | 54 |
| Tabla 4: Promedio y desviación estándar en cada clase de hidrofobicidad calculados para cada set de datos. | 55 |
| Tabla 5: Valores de promedio y desviación estándar para cada clase en las energías de captación calculadas para ambos set de datos. | 56 |
| Tabla 6: Indicadores de performance para el modelo de cada algoritmo en el set reducido..... | 58 |
| Tabla 7: Indicadores de performance para el modelo de cada algoritmo en el set completo. | 59 |
| Tabla 8: Distribución de ejemplos del set reducido balanceado..... | 61 |
| Tabla 9: Indicadores de performance al usar <i>smote</i> para balancear el set reducido. | 62 |
| Tabla 10: Distribución de ejemplos del set completo balanceado. | 63 |
| Tabla 11: Indicadores de performance al usar <i>smote</i> para balancear el set completo. | 63 |
| Tabla 12: Escalas de hidrofobicidad. | 91 |
| Tabla 13: Escala de energía de captación..... | 100 |
| Tabla 14: Valores de promedio y desviación estándar por clase para ambos set de datos..... | 105 |
| Tabla 15: Promedio y desviación estándar por clase calculados para ambos set de datos..... | 106 |
| Tabla 16: Valores del promedio y la desviación estándar en cada clase para el volumen en el volumen en ambos set de datos..... | 106 |
| Tabla 17: Promedio y desviación estándar de cada clase para cada torsión propuesta en ambos set de datos. | 107 |
| Tabla 18: Valores de promedio y desviación estándar en cada clase de composición aminoacídica para ambos set de datos. | 107 |

| | |
|---|-----|
| Tabla 19: Valores del promedio y desviación estándar de cada clase, asociados con los átomos por capa en cada set de datos. | 108 |
| Tabla 20: Valores promedio y de desviación estándar de cada clase calculados para los radios de sitio de ambos set de datos. | 108 |