

A person with curly hair, seen from behind, wearing a dark coat, stands with hands on hips, looking out over a misty, mountainous landscape. The scene is atmospheric and slightly desaturated.

EDICIÓN GENÉTICA: CUESTIONES ÉTICAS Y RETOS BIOINFORMÁTICOS

Hecho por:

José Manuel Sánchez Aquilué

Historia

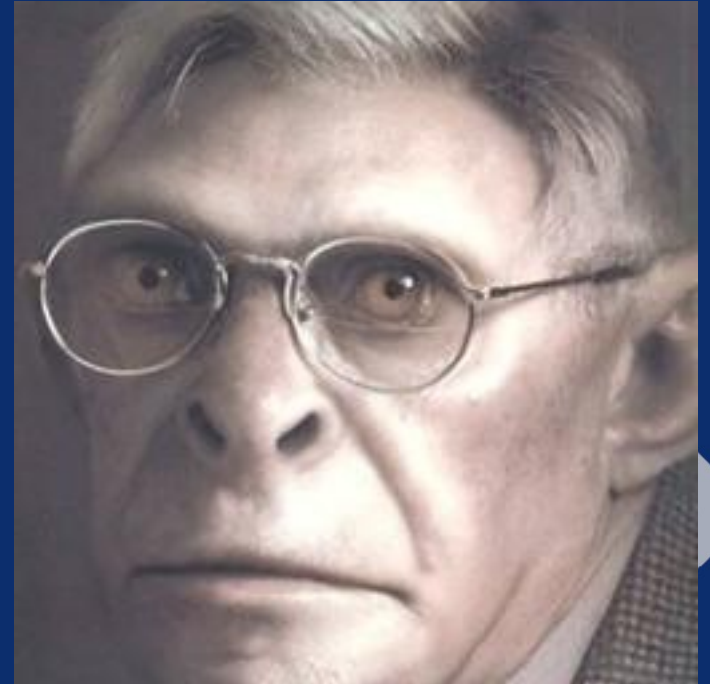
- Disciplina bastante reciente
- 1863 George Mendel sienta las bases de la genética
- 1953 Watson y Crick examinan la estructura del DNA
- 1970s primer DNA recombinado
- 1980s desarrollo de la PCR
- 2003 Proyecto Genoma Humano

Métodos de edición por nucleasas

- ZFPs/ZFNs (nucleasas de dedos de cinc). Coste de 30.000€. Principios de los 80
- TALEs/TALENs (nucleasa de actividad similar a activador de transcripción). Tiempo de 3 a 4 meses y 10.000€ de coste
- CRISPR-CAS. Tiempo de 2 a 3 semanas y costo de 20 a 30€.

Selección de genes

- Impulsa la evolución darwiniana
- Históricamente se ha practicado artificialmente con animales y plantas.
- PGD (Preimplantation Genetic Diagnosis)
- PGS (Preimplantation Genetic Screening)



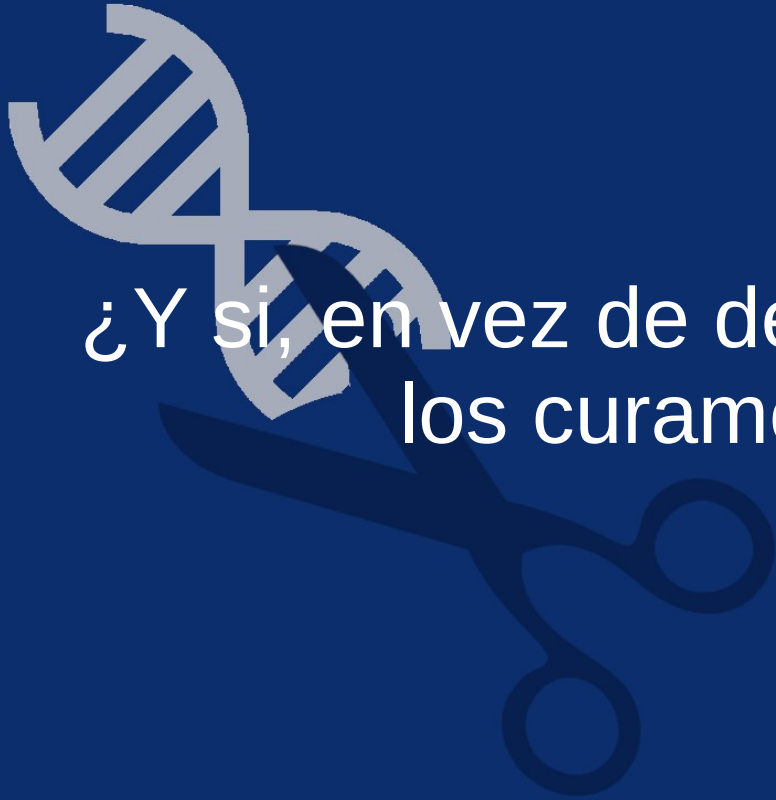
¿Qué dice la ley sobre la PGD?

- No puedes elegir genes a la carta. El donante es anónimo.
- Los embriones solo pueden ser rechazados si tienen una enfermedad genética rara de aparición temprana, incurable y potencialmente mortal
- Aquellos embriones rechazados se donan a otras parejas o a la investigación (existen límites)
- Gestación subrogada ilegal en España.

Críticas

- El embrión al contener ADN propio hay quien lo considera un ser humano.
- Seres humanos como medios en vez de fines en sí mismos.
- No es 100% efectiva, algunos embriones mueren
- El desarrollo de esta técnica podría vulnerar el derecho de intimidad de la persona

¿Y si, en vez de descartar embriones enfermos,
los curamos editando su DNA?

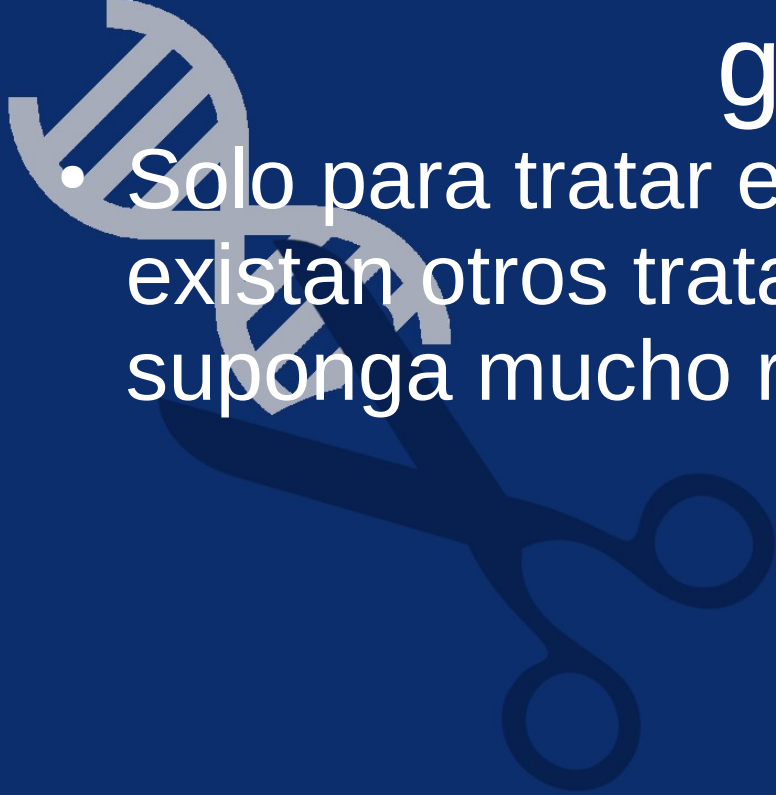


Problema off-target

- Se ha avanzado en anulación de genes
- Estamos lejos de editar características a nuestro antojo
- Primum non nocere
- Tenemos que elaborar una legislación que contemple los aspectos controvertidos de la edición genética

¿Qué dice la ley sobre la edición genética?

- Solo para tratar enfermedades y cuando no existan otros tratamientos o el tratamiento suponga mucho riesgo



No siempre ha sido así

- Principios del siglo XX, pérdida de poder de la religión y popularización del darwinismo social.
- Racismo muy presente en la sociedad occidental
- Leyes eugenésicas en Europa
- Tras la caída del Tercer Reich, se prohibieron estas prácticas



Consecuencias negativas de las “ingeniería perfecta”

- Aumento de la brecha social
- ¿Hasta donde puede cubrir la Seguridad Social?
- Estigma
- Explotación reproductiva y desigualdad de genero
- No todo es genética

Historia de He Jiankui

- Se saltó la ley para “ayudar” a un padre con VIH
- Editó el gen CCR5
- Una de las hijas no adquirió inmunidad
- Se piensa que podrían ser quiméricas
- CRISPS pudo afectar a otros genes
- CCR5 tiene otras funciones
- He Jiankui está condenado a cárcel



DAMOS PASO A LA
BIOINFORMÁTICA

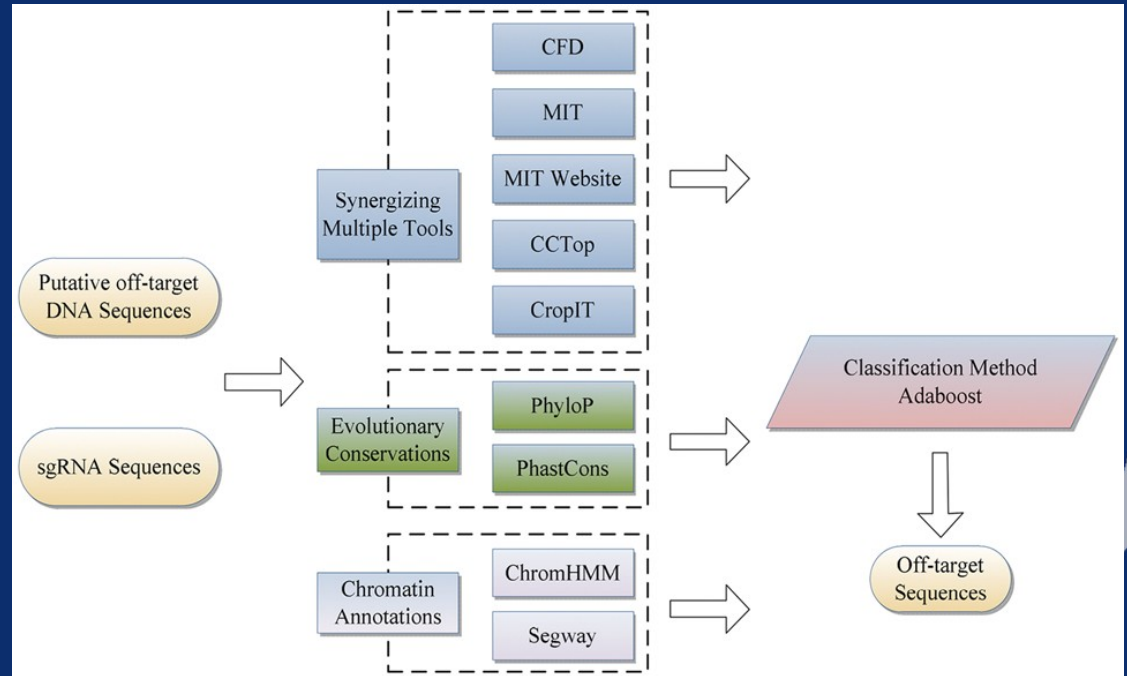


Técnicas de ML y DL para el problema off-target

- Experimento 1: técnicas clásicas de ML
(<https://github.com/Alexzsx/CRISPR>)
- Experimento 2: redes neuronales prealimentadas y convolucionales
(https://github.com/MichaelLinn/off_target_prediction)
- Experimento 3: más redes neuronales convolucionales
(<https://github.com/LQYoLH/CnnCrispr>)

Exp1. Input

- Dataset de 25332 secuencias, de las cuales 152 provocan off-target
- 3 grupos de predictores, un total de 9



Bioinformatics, Volume 35, Issue 7, 01 April 2019, Pages 1108–1115,
<https://doi.org/10.1093/bioinformatics/bty748>

The content of this slide may be subject to copyright: please see the slide notes for details.

Exp1. Predictores

Métodos de puntuación de off-target

- Puntuación CFD
- Puntuación CCTop
- Puntuación Cropit
- MIT de la web
- Puntuación MIT



Exp1. Predictores(2)

Conservación evolutiva

- puntuación PhyloP
- puntuación PhastCons

Anotaciones de cromatina

- ChromHMM
- Segway

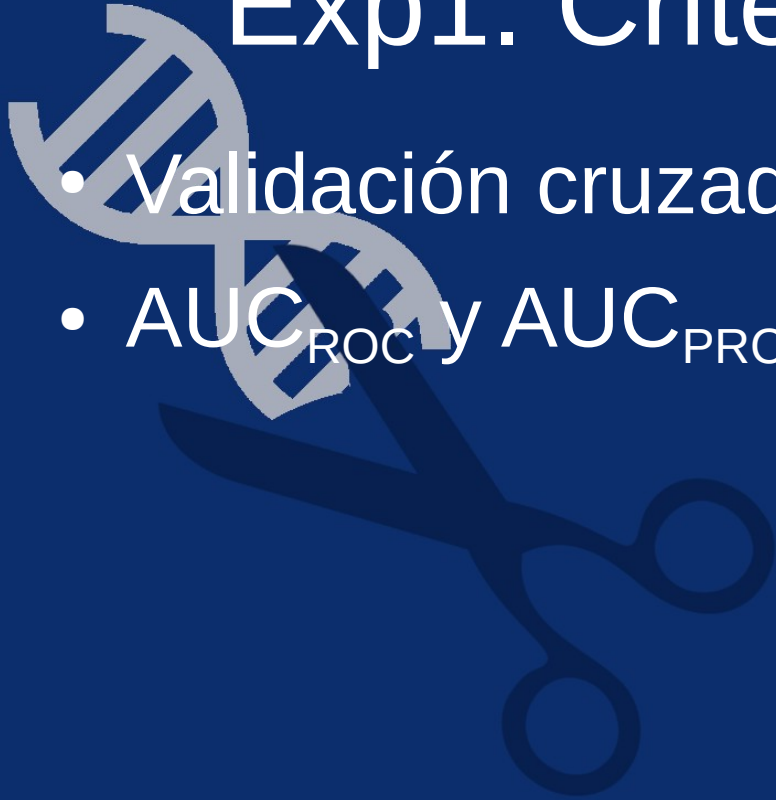
Exp1. Clasificadores

- AdaBoost
- Random Forest
- Multilayer Perceptron (MLP)
- Support Vector Machine (SVM)
- Árboles de decisión



Exp1. Criterios de evaluación

- Validación cruzada: 5-fold
- AUC_{ROC} y AUC_{PRC}



Exp1. Resultados

- Mejor los 5 métodos de puntuación juntos que por separado ($AUC_{ROC}=0.938, AUC_{PRC}=0.3$)
- PhyloP mejora la eficiencia pero PhastCons y las anotaciones cromáticas no
- El mejor método es AdaBoost y el peor SVM

Puntuación	AUC_{ROC}	AUC_{PRC}
Todos juntos	0.938	0.300
Puntuación CFD	0.915	0.139
Puntuación CCTop	0.765	0.108
Puntuación Cropit	0.806	0.113
MIT de la web	0.732	0.168
Puntuación MIT	0.872	0.192

Cuadro 4: Comparación entre los cinco métodos por separado y todos juntos.

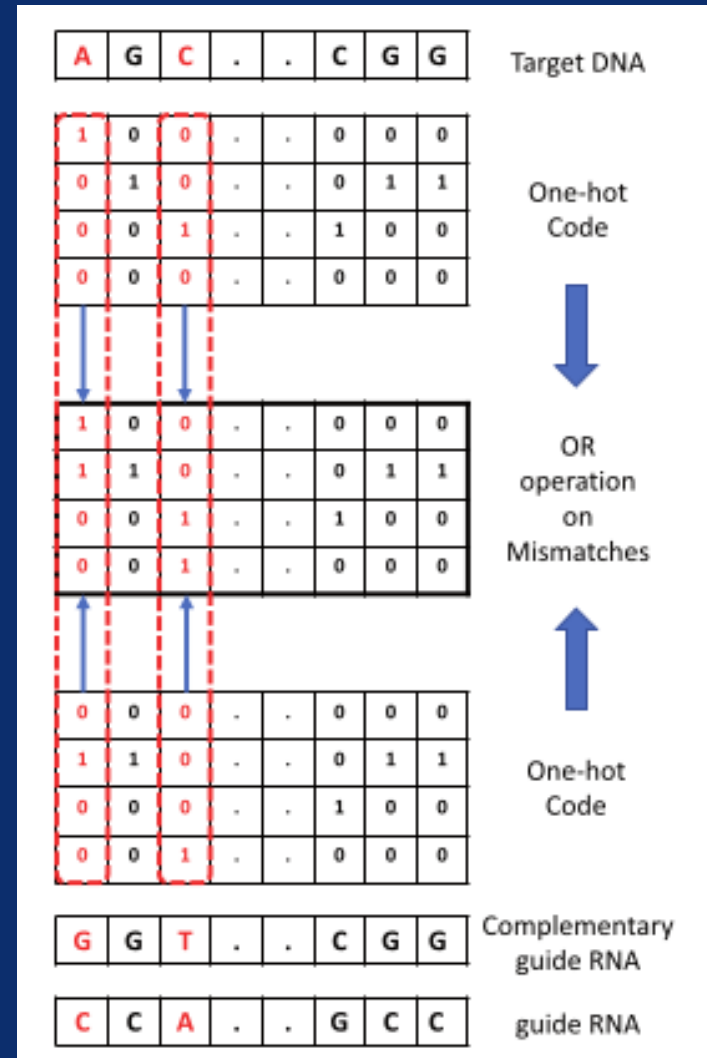
Puntuación	AUC_{ROC}	AUC_{PRC}
AdaBoost	0.9383	0.2998
Random Forest	0.8504	0.2236
Multilayer Perceptron (MLP)	0.9297	0.2431
Support Vector Machine (SVM)	0.6785	0.2038
Arboles de decisión	0.7981	0.2313

Cuadro 5: Comparación entre clasificadores.

Exp2. input

- Codificación one-hot de secuencias de 23 bases
- Combinar RNA guía y DNA
- 26034 secuencias, 143 offtarget
- 5 fold

Jiecong Lin, Ka-Chun Wong, Off-target predictions in CRISPR-Cas9 gene editing using deep learning, *Bioinformatics*, Volume 34, Issue 17, 01 September 2018, Pages i656–i663, <https://doi.org/10.1093/bioinformatics/bty554>



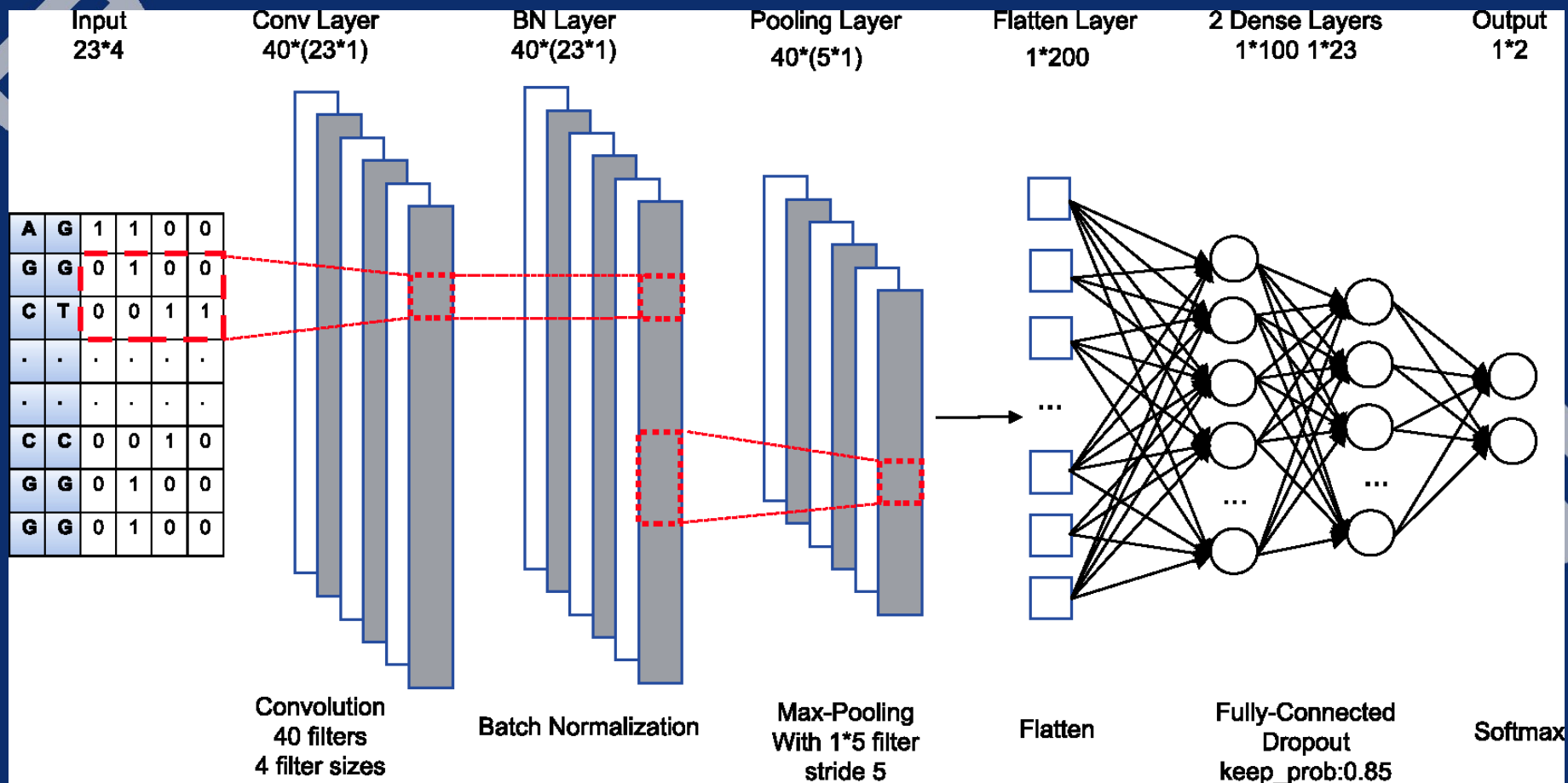
Exp2. Modelos

- Redes neuronales prealimentadas (FNN)
- Redes neuronales convolucionales (CNN)
- Comparativa con modelos clásicos de ML (LR, GBT, RF) y de puntuación (CFD, MIT, CCTop...)

Exp2. FNN

- Red neuronal no recurrente, no se forman ciclos.
- Con 2 capas ocultas: 250x40 neuronas
- Con 3 capas ocultas: 50x20x10 neuronas
- Con 4 capas ocultas: 25x10x10x4 neuronas

Exp2. CNN



Exp2. Variantes de CNN

- Estándar (figura anterior)
- Sin Normalización de Batch
- Sin drop-out
- Sin max-pooling
- Max-pooling con ventana de tamaño 3
- Max-pooling con ventana de tamaño 7

Exp2. resultados

- FNN_3capas y CNN estándar son los mejores modelos de NN
- BN y max-pooling son esenciales en CNN, drop-out no tanto
- Mejor que los modelos clásicos
- Mucho mejor que las puntuaciones

Modelo	AUC
FNN_2layer	0.842
FNN_3layer	0.970
FNN_4layer	0.954
CNN_std	0.972
CNN_nbn	0.954
CNN_nd	0.969
CNN_np	0.899
CNN_pool_win3	0.903
CNN_pool_win7	0.967

Cuadro 6: Comparación entre distintas variantes de redes neuronales.

Modelo	AUC	Modelo	AUC
FNN_3layer	0.970	FNN_3layer	0.970
CNN_std	0.972	CNN_std	0.972
Regresión Logística	0.931	puntuación CROP-IT	0.807
Gradient boosting	0.914	puntuación CFD	0.912
		puntuación MIT	0.865
		puntuación CCTop	0.776
		puntuación MIT-web	0.728

Cuadro 7: Comparación entre redes neuronales, modelos clásicos y puntuaciones.

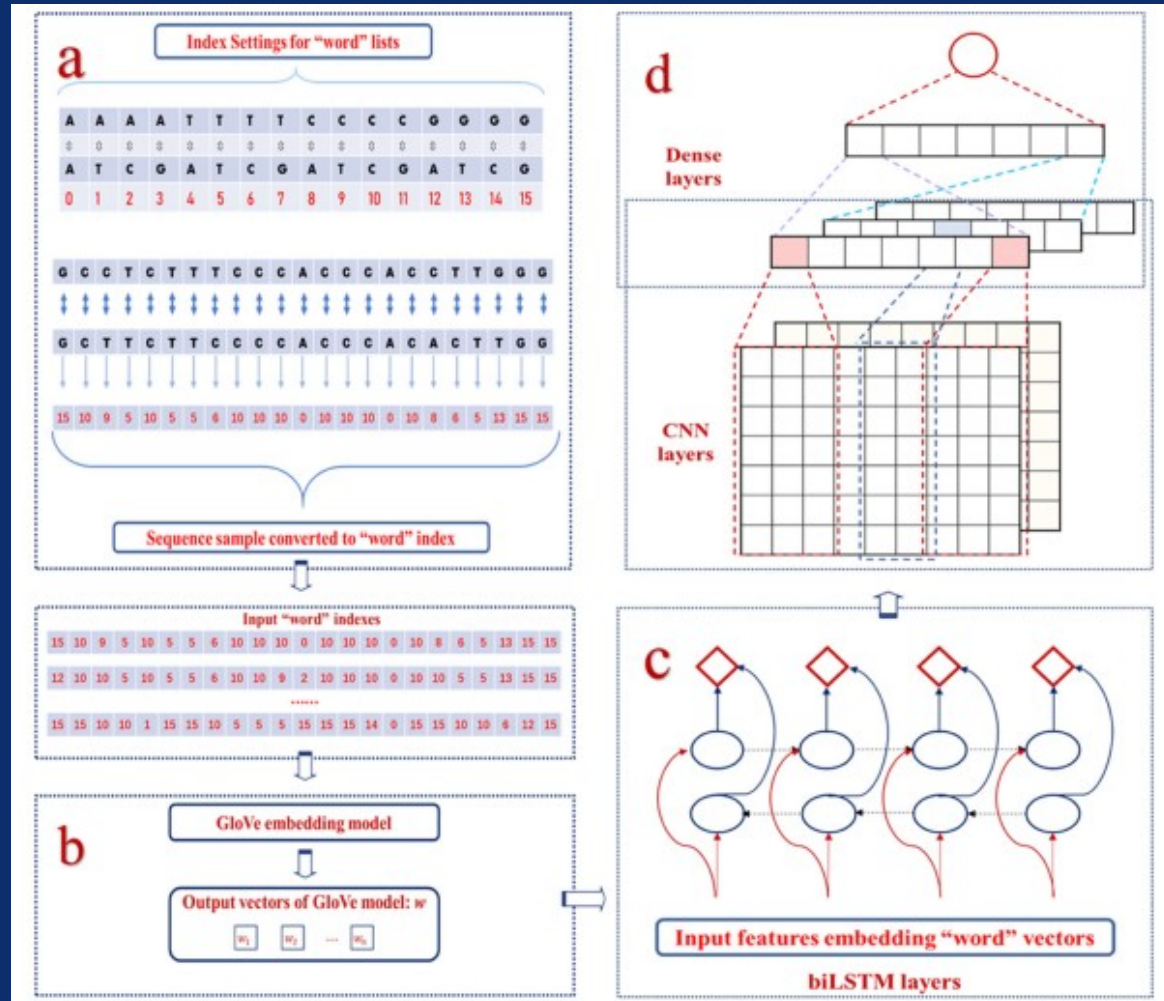
Exp3. Input

- Cada par de secuencias de sgRNA-DNA se codifica como un vector de 23 índices de palabras
- Cada combinación de dos bases en la misma posición se codifica como un índice
- Existen un total de 16 índices porque hay 4 nucleótidos

Exp.3 modelo

- Un modelo GloVe de embedding recibe los índices de palabras
- Tras el proceso de embedding obtenemos una matriz de 16×100 que sirve de input para una Long short-term memory (LSTM)
- El output de la (LSTM) alimenta 5 capas convolucionales.

- Batch normalization entre capas
- Finalmente se aplanan, se aplica un drop-out de 0.3 y le siguen dos capas densas de 20 y 2 neuronas.



Exp3. variantes

- CnnCrispr estándar
- Sin la LSTM bidireccional
- Invirtiendo el orden de las capas convolucional y la LSTM
- Sin batch normalization
- Sin drop-out

Exp3. resultados

- Sin BN no funciona en absoluto
- Sin drop-out mejora AUC, pero baja la exhaustividad
- Sin LSTM o cambiando sube AUC pero baja con creces la exhaustividad

Exp3. resultados

- El mejor tipo de modelo
- Una red neuronal convolucional profunda también da buenos resultados
- Con este data set la puntuación no funciona tan mal como en los experimentos anteriores

Modelo	Exhaustividad	AUC_{ROC}	AUC_{PRC}
CnnCrispr	0.857	0.975	0.679
CnnCrispr_NoLSTM	0.611	0.987	0.651
CnnCrispr_Conv_LSTM	0.643	0.986	0.67
CnnCrispr_NoBatchNor	-	0.5	0.504
CnnCrispr_NoDropOut	0.810	0.985	0.625

Cuadro 8: Comparación entre modelos de redes neuronales convolucionales.

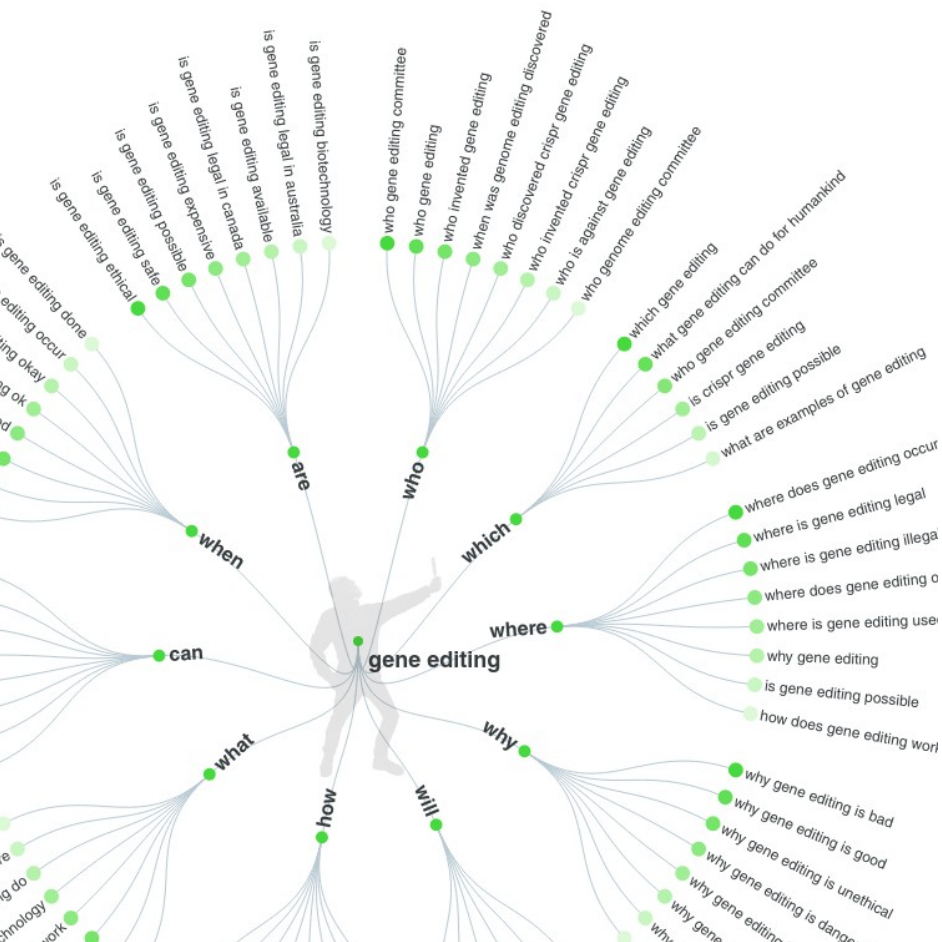
Modelo	AUC_{ROC}	AUC_{PRC}
CnnCrispr	0.975	0.679
CFD	0.942	0.316
MIT	0.77	0.044
CNN_std	0.947	0.208
DeepCrispr	0.981	0.497

Cuadro 9: Comparación entre el modelo de Qiaoyue Liu y otros métodos.



MUCHÍSIMAS
GRACIAS

Preguntas



Referencias

- Ley 14/2006, de 26 de mayo 2003, sobre técnicas de reproducción humana asistida. («BOE»núm. 126).
- Ley 35/1988, de 22 de noviembre 1988, sobre técnicas de reproducción asistida. («BOE» núm.282).
- Ley 45/2003, de 21 de noviembre 2003, por la que se modifica la ley 35/1988 sobre técnicasde reproducción asistida. («BOE» núm. 280).
- Ley orgánica 10/1995, de 23 de noviembre 1995, del código penal. («BOE» núm. 281).

Referencias

- Yan J Chen M Hong N Xue D Zhou C Zhu C Chen K Duan B Chuai G, Ma H. Deepcrispr: optimized crispr guide rna design by deep learning. *Genome Biol.*
- Adam P Cribbs and Sumeth M W Perera. Science and bioethics of crispr-cas9 gene editing: An analysis towards separating facts and fiction. *The Yale journal of biology and medicine*, 90(4):625–634, 2017.
- Rothschild J. Ethical considerations of gene editing and genetic selection. *Journal of general and family medicine*, 21:37–47, 2020.

Referencias

- Juan-Ramón Lacadena. Edición genómica: ciencia y ética. Revista Iberoamericana de Bioética,(3):1–16, jun. 2017.
- Jiecong Lin and Ka-Chun Wong. Off-target predictions in CRISPR-Cas9 gene editing using deep learning. Bioinformatics, 34(17):i656–i663, 09 2018.
- Cheng X. Liu G. et al Liu, Q. Deep learning improves the ability of sgrna off-target propensity prediction. BMC Bioinformatics, 21(51), 2020

Referencias

- Alexis Molina and Laura Serra. Genetic scissors for a dystopian future.ESCI news, 2021.
- Shixiong Zhang, Xiangtao Li, Qiuzhen Lin, and Ka-Chun Wong. Synergizing CRISPR/Cas9 off-target predictions for ensemble insights and practical applications.Bioinformatics,35(7):1108–1115, 08 2018