

UNIDAD DE DATASCIENCE

SINGLE CELL RNA AUTOENCODERS

Sebastián Cancino Espinoza

SUMMARY OF CONTENTS

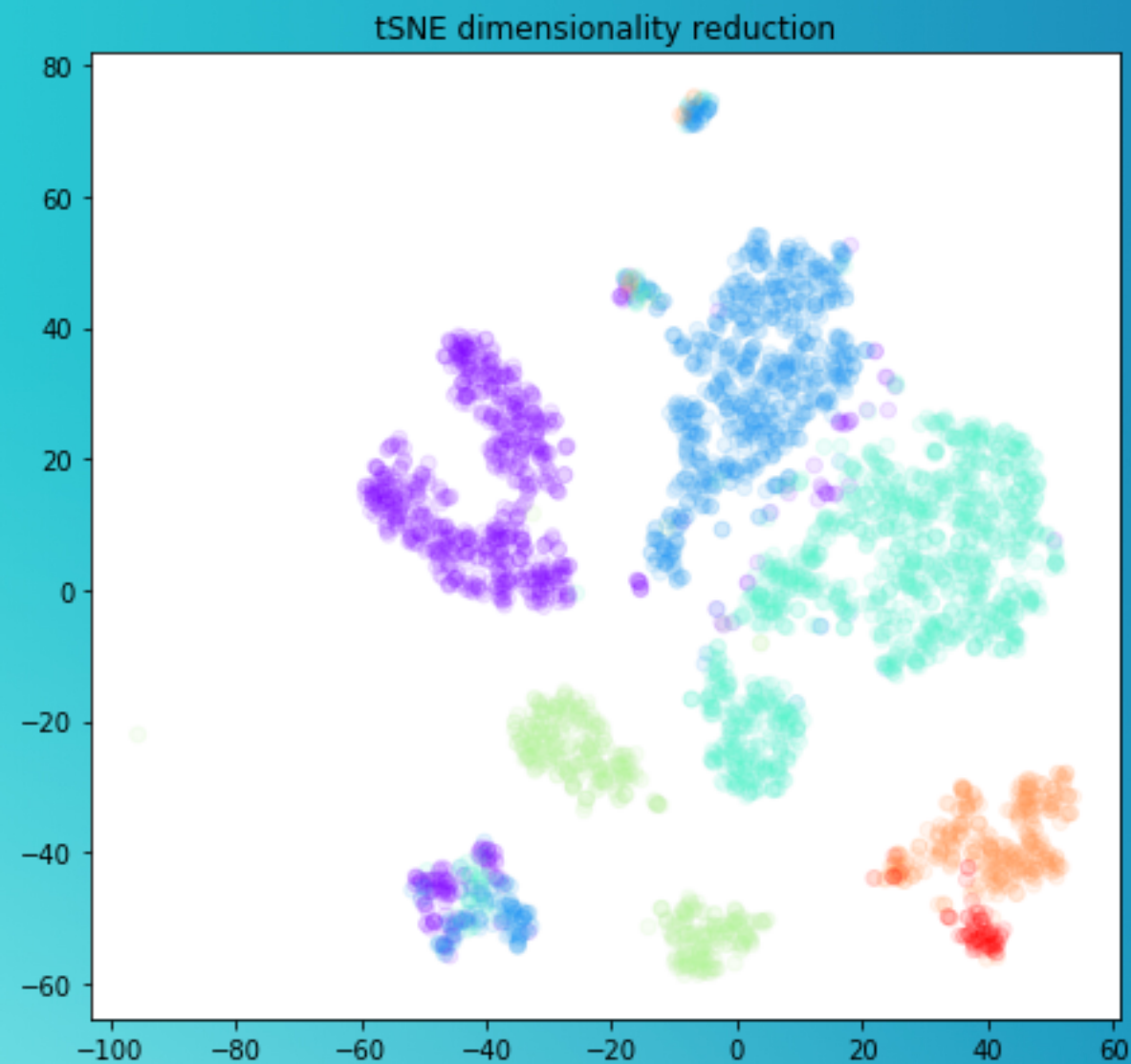
Single Cell RNA Seq
Literature Review
Data
Models
Preliminary Results
Discussion

WHAT IS SC-RNA-SEQ?

	Cell 1	Cell 2	...	Cell M
Gen 1	1	5	...	0
Gen 2	0	8	...	2
Gen 3	3	0	...	10
.
.
.
Gen N	2	1	...	4

EXPRESSION MATRIX (RAW COUNTS)

LITERATURE REVIEW



DIMENSIONALITY REDUCTION

- PCA
- tSNE
- ZIFA
- AutoEncoders
- VariationalAutoEncoders

DATASETS (GLIOMA)

	N° Cells	N° Genes	Format*
GSE132172	72	20.564	$\log_2(\text{TPM}/10 + 1)$
GSE131928	24133	24.556	$\log_2(\text{TPM}/10 + 1)$
GSE102130	2458	22.639	$\log_2(\text{TPM}/10 + 1)$
GSE89567	6342	22.659	$\log_2(\text{TPM}/10 + 1)$
GSE70630	3712	22.556	$\log_2(\text{TPM}/10 + 1)$
GSE84465	3590	23.460	Raw Counts
GSE103224	23795	60.725	Raw Counts

*TPM: Transcripts Per Million

64.102

(37.146)

Cells

5.491

Genes

DATA PREPROCESSING

1. **INPUT: LOG2-TRANSFORMED EXP MATRIX**
2. **FILTER NOT EXPRESSED GENES**
3. **MIN MAX SCALING**

MODELS

AUTOENCODER [1]

1. Random Projection (n=2048)
2. Dense Layer, LeakyReLU (n=512)
3. Dense Layer, Linear (n=16)
4. Dense Layer, LeakyReLU (n=512)
5. Output Layer, Sigmoid (n=2048)

VARIATIONAL AUTOENCODER [2]

1. Input Layer
2. Dropout Layer (rate=0.5)
3. Encoder Network, ReLU activation (n = [512, 128, 32])
4. Latent Sampling Layer (dim=2)
5. Decoder Network, ReLU activation (n = [32, 128, 512])
6. Output Layer, Sigmoid
7. Zero Inflated Layer

[1] Geddes, T., Kim, T., Nan, L. et al. Autoencoder-based cluster ensembles for single-cell RNA-seq data analysis. BMC Bioinformatics 20, 660 (2019). <https://doi.org/10.1186/s12859-019-3179-5>

[2] Wang DF, Gu J. VASC: dimension reduction and visualization of single-cell RNA-seq data by deep variational autoencoder. Genomics Proteomics Bioinformatics. ;16:320–31.

RESULTS

GSE70630

1. PCA
2. tSNE
3. UMAP
4. AutoEncoder
5. Variational AutoEncoder

DISCUSSION

1

EXPRESSION MATRIX FORMAT

TPM, raw counts

3

RECONSTRUCTIONS

2

SCALING METHOD

4

LATENT SPACE

UNIDAD DE DATASCIENCE

SINGLE CELL RNA AUTOENCODERS

Sebastián Cancino Espinoza