Multilayer perceptron

Utilizando los parámetros iniciales del experimento (3 epochs, batch_size = 128, embeddings freezados = True, etc), se probaron distintos algoritmos de optimización. Algunos son:

1. **Adadelta** (loss: CrossEntropyLoss):

▾ Notes

None

▾ Parameters

Name	Value
dropout	0.3
embeddings	./data/SBW-vectors-300-min5.txt.gz
embeddings_size	300
epochs	3
hidden_layers	[256, 128]
model_type	Multilayer Perceptron

▾ Metrics

Name	Value
train_loss	6.313
validation_bacc	0.002
validation_loss	6.313

2. **Adamax** (loss: CrossEntropyLoss):

▾ Notes

None

▾ Parameters

Name	Value
dropout	0.3
embeddings	./data/SBW-vectors-300-min5.txt.gz
embeddings_size	300
epochs	3
hidden_layers	[256, 128]
model_type	Multilayer Perceptron

▾ Metrics

Name	Value
train_loss	5.562
validation_bacc	0.041
validation_loss	5.486

3. **ASGD** (loss: CrossEntropyLoss):

▾ Notes

None

▾ Parameters

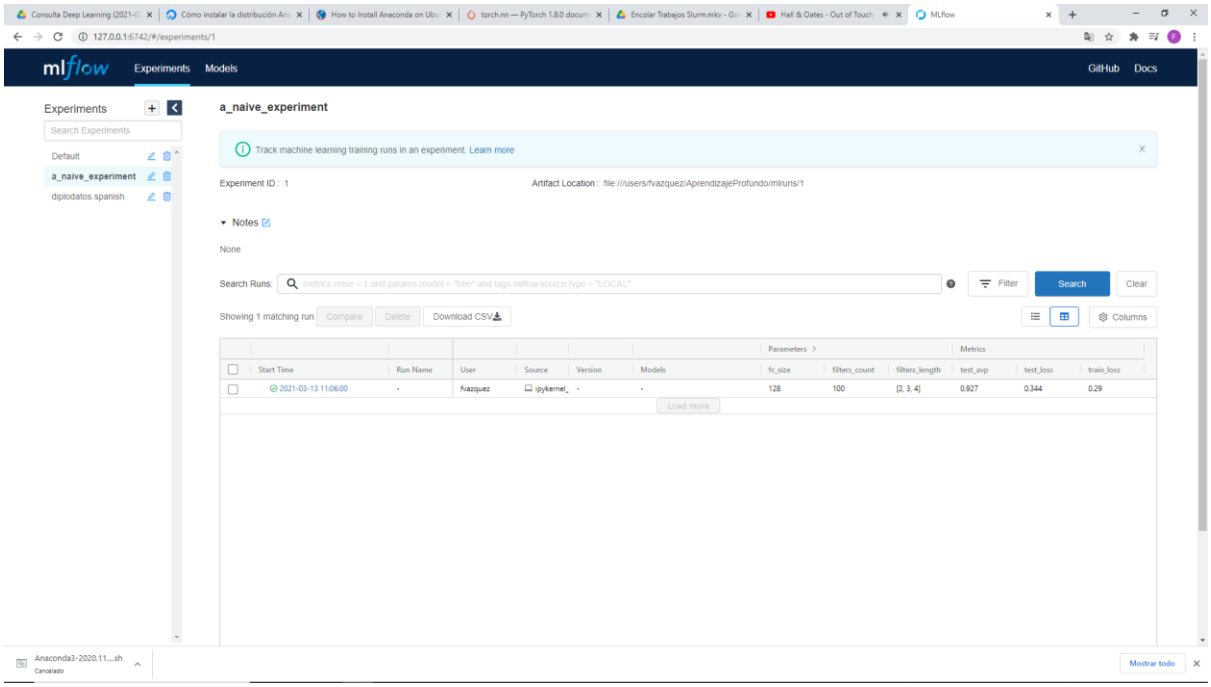
Name	Value
dropout	0.3
embeddings	./data/SBW-vectors-300-min5.txt.gz
embeddings_size	300
epochs	3
hidden_layers	[256, 128]
model_type	Multilayer Perceptron

▾ Metrics

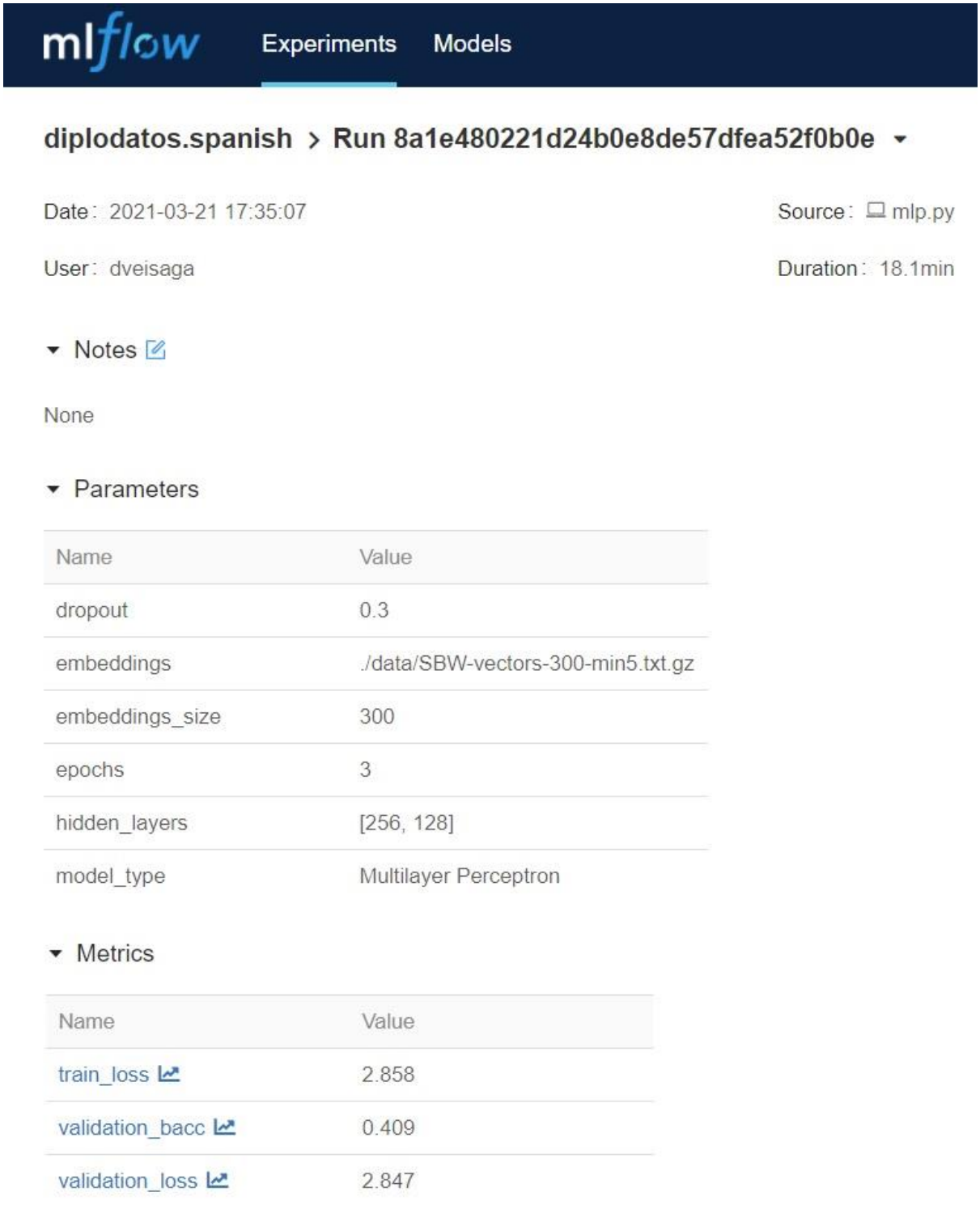
Name	Value
train_loss	6.445
validation_bacc	0.002
validation_loss	6.444

Nota: Lo que notamos (y no entendemos por qué) fue que probando distintas loss functions (BCEloss, GaussianNLLoss, MultilabelSoftMarginLoss) el modelos fallaba.

Mejor modelo:
Por otro lado, en uno de los primeros intentos logramos este resultado, que lamentablemente (y por inexperiencia) no lo pudimos trackear, pero sin duda fue el mejor resultado obtenido.



- **Corrida 349**
  `MLP Experimento con "batch_size"=64`



- A partir de Acá en cada captura de mlrun se encontrara el detalle de Hiperparametros Modificados

diplodatos.spanish > Run 8a8d9236a2da45399e77f431a404dccb

Date: 2021-03-21 20:59:12

Duration: 12.0min

▾ Notes

mlp5.py
Loss: NLLLoss

▾ Parameters

Name	Value
dropout	0.3
embeddings	./data/SBW-vectors-300-min5.txt.gz
embeddings_size	300
epochs	3
hidden_layers	[256, 128]
model_type	Multilayer Perceptron

▾ Metrics

Name	Value
train_loss	-54208957432
validation_bacc	0.002
validation_loss	-88007890746.5

diplodatos.spanish > Run 1a84a9b5dfdb4016b873918af99c9bcc

Date: 2021-03-21 17:53:15

Duration: 17.1min

Notes

mlp4.py
Se modifico lo siguiente

- **Optimizador: Adagrad(lr = 1e-3, lr_decay= 0.001, weigth_decay= 1e-5. eps = 1e-10)**
- **freeze embeddings** = False
- **random_buffer_size = 2200**

Parameters

Name	Value
dropout	0.3
embeddings	./data/SBW-vectors-300-min5.txt.gz
embeddings_size	300
epochs	3
hidden_layers	[256, 128]
model_type	Multilayer Perceptron

Metrics

Name	Value
train_loss	2.918
validation_bacc	0.295
validation_loss	2.908

diplodatos.spanish > Run 01352524a07f4e179f42408fa5a7e2e4

Date: 2021-03-21 16:59:07

Duration: 20.9min

Notes

mlp3.py
En esta corrida se utilizo:

- **Optimizador: SGD con lr= 0.2 y momentum = 0.9**
- **Epochs: 6**

Parameters

Name	Value
dropout	0.3
embeddings	./data/SBW-vectors-300-min5.txt.gz
embeddings_size	300
epochs	6
hidden_layers	[256, 128]
model_type	Multilayer Perceptron

Metrics

Name	Value
train_loss	2.869
validation_bacc	0.402
validation_loss	2.854

diplodatos.spanish > Run 655122808e5d4185b38e71b78b60cb4b

Date: 2021-03-21 16:34:26 Sour

Duration: 11.1min Statu

Notes

mlp2.py
Se cambio:

Optimizador: ASGD (lr=1e-2, lambd = 0.0001, alpha = 0.75, t0 = 1e6, weigth decay = 1e-4)

Parameters

Name	Value
dropout	0.3
embeddings	./data/SBW-vectors-300-min5.txt.gz
embeddings_size	300
epochs	3
hidden_layers	[256, 128]
model_type	Multilayer Perceptron

Metrics

Name	Value
train_loss	5.974
validation_bacc	0.027
validation_loss	5.793

diplodatos.spanish > Run 0508d8d86f444bcd852944188b31e05f

Date: 2021-03-20 20:25:27

Duration: 13.1min

Notes

mlp1.py

- **optimizador: ASGD (lr=1e-4, alpha=0.5, weight_decay=1e-3)**
- **loss: BCEWithLogitsLoss**

Parameters

Name	Value
dropout	0.3
embeddings	./data/SBW-vectors-300-min5.txt.gz
embeddings_size	300
epochs	3
hidden_layers	[256, 128]
model_type	Multilayer Perceptron

Metrics

Name	Value
train_loss	2.846
validation_bacc	0.413
validation_loss	2.821