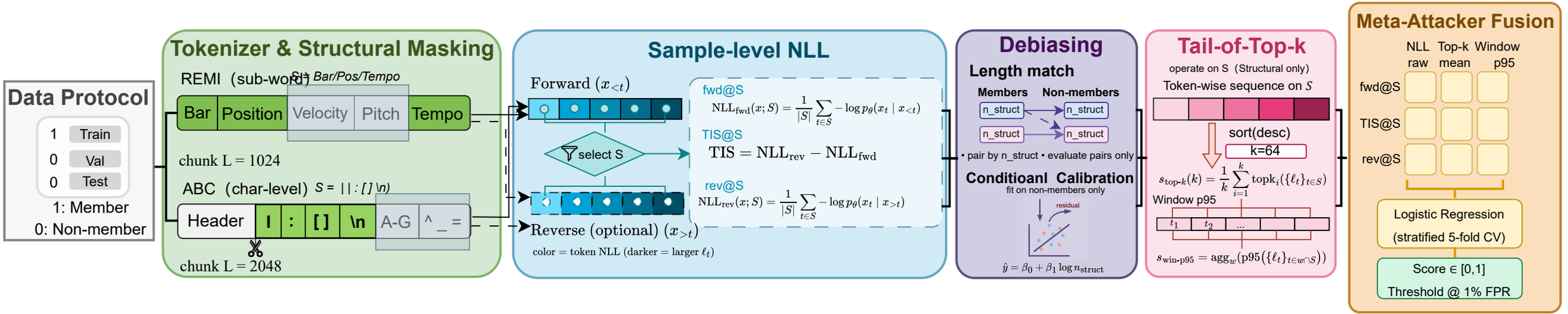
Data Protocol
1 Train
0 Val
0 Test
1: Member
0: Non-member
Tokenizer & Structural Masking
REMI (sub-word)
Bar/Pos/Tempo
Bar
Position
Velocity
Pitch
Tempo
chunk L = 1024
ABC (char-level)
S = | | : [ ] \n)
Header
|
:
[ ]
\n
A-G
^ _ =
chunk L = 2048
Sample-level NLL
Forward ($x_{<t}$)
select S
Reverse (optional) ($x_{>t}$)
color = token NLL (darker = larger ℓ_t)
fwd@S
$\mathrm{NLL}_{\mathrm{fwd}}(x;S) = \frac{1}{|S|}\sum_{t\in S} -\log p_\theta(x_t \mid x_{<t})$
TIS@S
$\mathrm{TIS} = \mathrm{NLL}_{\mathrm{rev}} - \mathrm{NLL}_{\mathrm{fwd}}$
rev@S
$\mathrm{NLL}_{\mathrm{rev}}(x;S) = \frac{1}{|S|}\sum_{t\in S} -\log p_\theta(x_t \mid x_{>t})$
Debiasing
Length match
Members
Non-members
n_struct
n_struct
n_struct
n_struct
• pair by n_struct • evaluate pairs only
Conditioanl Calibration
fit on non-members only
residual
$\hat{y} = \beta_0 + \beta_1 \log n_{\mathrm{struct}}$
Tail-of-Top-k
operate on S (Structural only)
Token-wise sequence on S
sort(desc)
k=64
$s_{\text{top-}k}(k) = \frac{1}{k}\sum_{i=1}^{k} \mathrm{topk}_i(\{\ell_t\}_{t\in S})$
Window p95
t_1
t_2
...
$s_{\text{win-p95}} = \mathrm{agg}_w(\mathrm{p95}(\{\ell_t\}_{t\in w\cap S}))$
Meta-Attacker Fusion
NLL raw
Top-k mean
Window p95
fwd@S
TIS@S
rev@S
Logistic Regression
(stratified 5-fold CV)
Score ∈ [0,1]
Threshold @ 1% FPR