

Data Model of Generate Learning Sets



SEGREGATION SYSTEM CONFIGURATION	
+ configurationID	: UUID
+ preparationIP	: String
+ messagingIP	: String
+ preparationIP	: String
+ requestedSessions	: Integer
+ sufficientSessionsThreshold	: Integer
+ maxim imbalanceTolerance	: float
+ trainingSetPercentage	: float
+ validationSetPercentage	: float
+ testSetPercentage	: float

<enumeration>	
PARTITION TYPE	
TRAINING	
VALIDATION	
TEST	

PREPARED SESSION	
+ sessionID	: UUID
+ rawSessionID	: UUID
+ speakerID	: UUID
+ featureVectorID	: UUID
+ processedAt	: DateTime
+ missingSamples	: Integer
+ maxMissingSamples	: Integer
+ status	: PreparationStatus
+ label	: Label

LEARNING SET	
+ learningSetID	: UUID
+ generatedAt	: DateTime
+ configurationID	: UUID
+ totalSessions	: Integer
+ trainingSessions	: Integer
+ validationSessions	: Integer
+ testSessions	: Integer

LEARNING SET ITEM	
+ itemID	: UUID
+ learningSetID	: UUID
+ preparedSessionID	: UUID
+ type	: PartitionType

DATA BALANCING REPORT	
+ reportID	: UUID
+ generatedAt	: DateTime
+ configurationID	: UUID
+ totalSessions	: Integer
+ angrySessions	: Integer
+ happySessions	: Integer
+ neutralSessions	: Integer
+ sadSessions	: Integer
+ surpriseSessions	: Integer

FEATURE COVERAGE REPORT	
+ reportID	: UUID
+ generatedAt	: DateTime
+ featureName	: String
+ minValue	: Double
+ maxValue	: Double
// Possible fields:	
+ meanValue	: Double
+ stdValue	: Double
+ coveredRatio	: Double