

1: BUSINESS UNDERSTANDING		2: DATA UNDERSTANDING		3: DATA PREPARATION	4: MODELING
GENERIC	Business goals/shortfalls	Data collection		Data table & records	Model applicability
	Performance areas	Data quality assessment		Data attributes	Model selection
	Success/failure metrics	Data subset & clustering		Data standardization	Model output
Client (Standard)	<ul style="list-style-type: none">Overall underwriting performance	<ul style="list-style-type: none">Internal data / 240 pointsCredit applicationsInternal related data3rd party related data		<ul style="list-style-type: none">Manual/code based	<ul style="list-style-type: none">Credit/qualitativeCredit/Scoring & QuantRegression based
Client (Other Labs)	<ul style="list-style-type: none">Other Labs' subset goals	<ul style="list-style-type: none">Other Labs' alternative data		<ul style="list-style-type: none">Other Labs' data prep methods	<ul style="list-style-type: none">Other Labs' AI/ML models
Client Proposal	Specific subsets/goals: <ul style="list-style-type: none">First 12 months/no history for credit riskSME /limited data & analyticsFraud + mismanagementIncrease underwriting connectivity <i>Parallel time line see next table.</i> <i>Priced separately in proposal.</i>	New/add. alternative data: <ul style="list-style-type: none">Social media & big data<ul style="list-style-type: none">Web site postings from companySocial media posting by companyDigitalized news about companyWhen applicable: Changes in stock/credit ratingsSatellite imagesCredit card receiptsEmail inboxesDrone imagesIoTAdditional based on poof-of-concept findings		New/add. data prep methods: <ul style="list-style-type: none">Data endogenizationKernel methodsBasic expansion & regularization	New/add. modeling techniques: <ul style="list-style-type: none">Linear/non-linear regressionSupervised learning<ul style="list-style-type: none">Bayesian models & inferenceAdditive, trees, and relatedNeural NetworksPrototype & nearest neighborsFine grained proximity analysisUnsupervised learning<ul style="list-style-type: none">Random forest & ensembleGraphical modelsHigh dimensionalAdditional based on proof-of-concept

LEGEND
3-month proof-of-concept
12-month scope
Existing / other Labs

5: EVALUATION
Business objectives
Success/failure metrics
Business value
Key issues overlooked
Decisions/Recs on data selection
Decisions/Recs on model selection

6: DEPLOYMENT
Technology selection
Implementing repeatable scoring
Standardize data-selection
Standardize data input selection
Standardize data structuring
Standardize modeling

1. BUSINESS UNDERSTANDING		2. DATA UNDERSTANDING		3. DATA PREPARATION		4. MODELING		5. EVALUATION		6. DEPLOYMENT		
3-MONTH PROOF-OF-CONCEPT	DATA MODELING	PROJECT INCEPTION		DATA								
		WEEK 1		WEEK 2		WEEK 3		WEEK 4		WEEK 5		
		Meetings wt. internal stakeholders		Standardized	Creating data list/request Collection, assessment, subset, clustering		Table, records, attributes, standardization Endogenization, kernel, expansion, reg.		Linear/non-linear regression Supervised learning 			