// This file contains messages for various machine learning inferences

// such as regression and classification.

//

// In many applications more than one type of inference is desired for a single

// input. For example, given meteorologic data an application may want to

// perform a classification to determine if we should expect rain, snow or sun

// and also perform a regression to predict the temperature.

// Sharing the single input data between two inference tasks can be accomplished

// using MultiInferenceRequest and MultiInferenceResponse.

syntax = "proto3";

option cc\_enable\_arenas = true;

import "tensorflow\_serving/apis/classification.proto";

import "tensorflow\_serving/apis/input.proto";

import "tensorflow\_serving/apis/model.proto";

import "tensorflow\_serving/apis/regression.proto";

package tensorflow.serving;

// Inference request such as classification, regression, etc...

message InferenceTask {

// Model Specification. If version is not specified, will use the latest

// (numerical) version.

// All ModelSpecs in a MultiInferenceRequest must access the same model name.

ModelSpec model\_spec = 1;

// Signature's method\_name. Should be one of the method names defined in

// third\_party/tensorflow/python/saved\_model/signature\_constants.py.

// e.g. "tensorflow/serving/classify".

string method\_name = 2;

}

// Inference result, matches the type of request or is an error.

message InferenceResult {

ModelSpec model\_spec = 1;

oneof result {

ClassificationResult classification\_result = 2;

RegressionResult regression\_result = 3;

}

}

// Inference request containing one or more requests.

message MultiInferenceRequest {

// Inference tasks.

repeated InferenceTask tasks = 1;

// Input data.

Input input = 2;

}

// Inference request containing one or more responses.

message MultiInferenceResponse {

// List of results; one for each InferenceTask in the request, returned in the

// same order as the request.

repeated InferenceResult results = 1;

}