

Apache MXNet (incubating) 1.1.0 Release Notes

- Usability Improvements
- Bug-fixes
- New Features
- API Changes
- Deprecations
- Performance Improvements
- Known Issues
- How to build MXNet
 - List of submodules used by Apache MXNet (Incubating) and when they were updated last

Usability Improvements

- Improved the usability of examples and tutorials
- C++ Stack traces are now demangled in OSX and Linux (improved developer experience and bug reports).

Bug-fixes

- Fixed I/O multiprocessing for too many open file handles (#8904), race condition (#8995), deadlock (#9126).
- Fixed image IO integration with OpenCV 3.3 (#8757).
- Fixed Gluon block printing (#8956).
- Fixed float16 argmax when there is negative input. (#9149)
- Fixed random number generator to ensure sufficient randomness. (#9119, #9256, #9300)
- Fixed custom op multi-GPU scaling (#9283)
- Fixed gradient of gather_nd when duplicate entries exist in index. (#9200)
- Fixed overridden contexts in Module `group2ctx` option when using multiple contexts (#8867)
- Fixed `swap_axes` operator with "add_to" gradient req (#9541)
- Fixed SSD kernel launch resource limits on volta (#9480)

New Features

- Added experimental API in `contrib.text` for building vocabulary, and loading pre-trained word embeddings, with built-in support for 307 GloVe and FastText pre-trained embeddings. (#8763)
- Added experimental structural blocks in `gluon.contrib`: `Concurrent`, `HybridConcurrent`, `Identity`. (#9427)
- Added `sparse.dot(dense, csr)` operator (#8938)
- Added `Khatri-Rao` operator (#7781)
- Added `FTML` and `Signum` optimizer (#9220, #9262)
- Added `ENABLE_CUDA_RTC` build option (#9428)

API Changes

- Added zero gradients to rounding operators including `rint`, `ceil`, `floor`, `trunc`, and `fix` (#9040)
- Added `use_global_stats` in `nn.BatchNorm` (#9420)
- Added `axis` argument to `SequenceLast`, `SequenceMask` and `SequenceReverse` operators (#9306)
- Added `lazy_update` option for standard `SGD` & `Adam` optimizer with `row_sparse` gradients (#9468, #9189)
- Added `select` option in `Block.collect_params` to support regex (#9348)
- Added support for (one-to-one and sequence-to-one) inference on explicit unrolled RNN models in R (#9022)

Deprecations

- The Scala API name space is still called "ml.dmlc". The name space is likely be changed in a future release to "org.apache" and might brake existing applications and scripts (#9579, #9324)

Performance Improvements

- Improved GPU inference speed by 20% when batch size is 1 (#9055)
- Improved `SequenceLast` operator speed (#9306)
- Added multithreading for the class of broadcast_reduce operators on CPU (#9444)
- Improved batching for GEMM/TRSM operators with large matrices on GPU (#8846)

Known Issues

- "Predict with pre-trained models" tutorial is broken
- "example/numpy-ops/ndarray_softmax.py" is broken

How to build MXNet

Please follow the instructions at <https://mxnet.incubator.apache.org/install/index.html>

List of submodules used by Apache MXNet (Incubating) and when they were updated last

Submodule:: Last updated by MXNet:: Last update in submodule

1. cub@: Jul 31, 2017 :: Jul 31, 2017
2. dlpack@: May 20, 2017 :: Oct 30, 2017
3. dmlc-core@: Jan 16, 2018 :: Jan 17, 2018
4. mshadow@: Jan 11, 2017 :: Jan 15, 2018
5. nnvm@: Dec 9, 2017 :: Jan 10, 2017
6. ps-lite@: Nov 21, 2017 :: Jan 2, 2017