Neural BLACKLIST

Aug 2017

Sean Park

Senior Malware Scientist ,Trend Micro

spark@trendmicro.com

HITCON

In The Old Days

Today

Filter by threat: Botnet C&Cs | Payment Sites | Distribution Sites

Filter by malware: TeslaCrypt | CryptoWall | TorrentLocker | PadCrypt | Locky | CTB-Locker | FAKBEN | PayCrypt | DMALocker | Cerber | Sage

Dateadded (UTC)	Threat	Malware	Host (?)	Domain Registrar (?)	IP address (ASN, Cou
2017-08-20 06:45	Payment Site	Cerber	qfjhpgbefuhenjp7.1e1jbc.top	Eranet International Limited	92.63.91.45 (Latvia)
2017-08-14 16:17	Payment Site	Cerber	oqwygprskqv65j72.1fs9pz.top	Eranet International Limited	104.244.156.10 (Un
2017-08-12 15:43	Payment Site	Cerber	oqwygprskqv65j72.14jqyo.top	Eranet International Limited	103.11.65.175 (Unit
2017-08-08 14:01	Payment Site	Cerber	oqwygprskqv65j72.1kh9ct.top	Eranet International Limited	103.11.65.175 (Unit
2017-08-04 10:52	Payment Site	Cerber	oqwygprskqv65j72.13rdvu.top	Eranet International Limited	103.11.65.165 (Unit
2017-07-31 18:19	Payment Site	Cerber	oqwygprskqv65j72.1hbdbx.top	Eranet International Limited	103.11.65.165 (Unit
2017-07-30 22:35	Payment Site	Cerber	oqwygprskqv65j72.13gpqd.top	Eranet International Limited	103.11.65.165 (Unit
2017-07-27 18:46	Payment Site	Cerber	qfjhpgbefuhenjp7.16g9ub.top	Eranet International Limited	107.181.161.207 (U
2017-07-25 14:58	Payment Site	Cerber	hjhqmbxyinislkkt.1jmip6.top	Eranet International Limited	155.94.213.132 (Un
2017-07-23 13:11	Payment Site	Cerber	qfjhpgbefuhenjp7.13iuvw.top	Eranet International Limited	107.181.161.207 (U
2017-07-21 01:20	Payment Site	Cerber	xpcx6erilkjced3j.1n5mod.top	Eranet International Limited	185.101.218.131 (U
2017-07-21 00:57	Payment Site	Cerber	qfjhpgbefuhenjp7.158ugp.top	Eranet International Limited	107.181.161.207 (U
2017-07-18 17:24	Payment Site	Cerber	hjhqmbxyinislkkt.1bcnad.top	Eranet International Limited	104.200.67.22 (Unit
2017-07-18 11:29	Payment Site	Cerber	qfjhpgbefuhenjp7.1fcfjn.top	Eranet International Limited	107.181.161.207 (U
2017-07-18 10:37	Payment Site	Cerber	xpcx6erilkjced3j.19kdeh.top	Eranet International Limited	107.150.18.186 (Un
2017-07-17 00:29	Payment Site	Cerber	hjhqmbxyinislkkt.18zrup.top	Eranet International Limited	104.200.67.22 (Unit
2017-07-14 22:00	Payment Site	Cerber	qfjhpgbefuhenjp7.1225wj.top	Eranet International Limited	107.181.161.207 (U

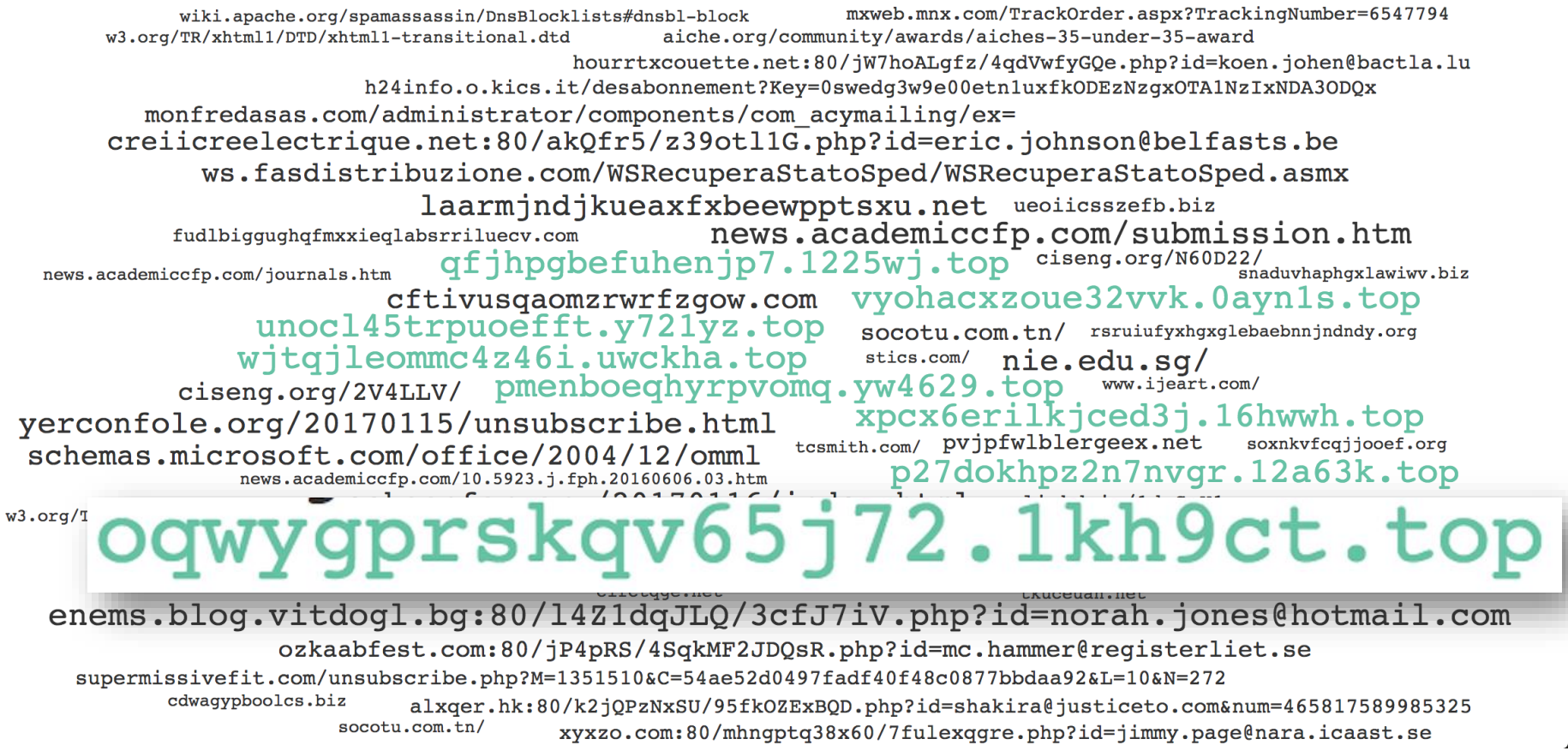
Source: https://ransomwaretracker.abuse.ch/tracker/cerber/

CryptoLocker

Locky


enems.blog.vitdogl.bg:80/l4Z1dqJLQ/3cfJ7iV.php?id=norah.jones@hotmail.com
loyabc.com:80/vfmc1mjskrtu/1apitn.php?id=kevin.baker@hbfutlers.com
h24info.o.kics.it/desabonnement?Key=0swedg3w9e00etn1uxfkODEzNzgxOTA1NzIxNDA3ODQx
mxweb.mnx.com/TrackOrder.aspx?TrackingNumber=6547794
monfredasas.com/administrator/components/com_acymailing/ex=
dokadfa.ru:80/mFovYD0C6pf/n1SH2k.php?id=peter.jackson@reynoraad.com
creiicreelectrique.net:80/akQfr5/z39otl1G.php?id=eric.johnson@belfasts.be
qfjhpgbefuhenjp7.1225wj.top
tcsmith.com/
snaduvhaphgxlawiwv.biz
wiki.apache.org/spamassassin/DnsBlocklists#dnsbl-block
laarmjndjkueaxfxbeewpptsxu.net
rsruiufyxhgxglebaebnnjndndy.org
w3.org/TR/html4/strict.dtd
peakconfor.org/20170116/index.html
clfctqge.net
nie.edu.sg/
vyohacxzoue32vvk.0ayn1s.top
fudlbiggughqfmxxieqlabsrriluecv.com
tkuceuah.net
socotu.com.tn/
pvjpfwlblergeex.net
schemas.microsoft.com/office/2004/12/omml
socotu.com.tn/
xpcx6erilkjced3j.16hwwh.top
www.ijeart.com/
unocl45trpuoefft.y721yz.top
awpjugdubjlpwmpqsfjb.com
pmenboeqnyrpvomq.yw4629.top
stics.com/
yerconfole.org/20170115/unsubscribe.html
soxnkvfcqjjooef.org
ciseng.org/N60D22/
hourrtxcouette.net:80/jW7hoALgfz/4qdVwfyGQe.php?id=koen.johen@bactla.lu
chdiekgopartylines.com:80/XgnGV4/4YaUQyNdXA.php?id=mobile@fastestgeb.it
oqwygprskqv65j72.1kh9ct.top
news.academiccfp.com/journals.htm
wjtqjleommc4z46i.uwckha.top
news.academiccfp.com/10.5923.j.fph.20160606.03.htm
p27dokhpz2n7nvgr.12a63k.top
ws.fasdistribuzione.com/WSRecuperaStatoSped/WSRecuperaStatoSped.asmx
w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd
xyxzo.com:80/mhngptq38x60/7fulexqgre.php?id=jimmy.page@nara.icaast.se
blog.interiextfilecasar.com:80/Z8iGogy/92oqcMLUphaeZv8.php?id=christopher.nolan@zoeaacs.com
news.academiccfp.com/10.5923.j.cmaterials.20160606.03.htm

RIGEK-Cerber



Problems


click.e.vineyardvines.com/?qs=6ba2b95d7aec369a2e70c9a04d61348a7684b633ffe55dc996a089b89c628b325075780cc9268cf4
news.academiccfp.com/10.5923.j.fph.20160606.03.htm
navi.mail.carenet.com/c0/m1/teid-t9YQTYIV8TB62mDtu2Owc7LmIboTQU1UdeSSV4/www.carenet.com/news/general/carenet/43305
monfredasas.com/administrator/components/com_acymailing/ex=
unocl45trpuoefft.y721yz.top
cruise.co.uk/news/?utm_campaign=NEWS220117&utm_medium=email&utm_source=NEWS220117&email=sh.smith@sdv.com
supermissivefit.com/unsubscribe.php?M=1351510&C=54ae52d0497fadf40f48c0877bbdaa92&L=10&N=272
cftivusqaomzrwrfzgow.com
news.academiccfp.com/journals.htm
w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd
enems.blog.vitdogl.bg:80/l4Z1dqJLQ/3cfJ7iV.php?id=norah.jones@hotmail.com
fudlbiggughqfmxxieqlabsrriluecv.com
p27dokhpz2n7nvgr.12a63k.top
wiki.apache.org/spamassassin/DnsBlocklists#dnsbl-block
news.academiccfp.com/10.5923.j.cmaterials.20160606.03.htm
ciseng.org/N60D22/
snaduvhaphgxlawiwv.biz
mxweb.mnx.com/TrackOrder.aspx?TrackingNumber=6547794
soxnkvfcqjjooef.org
rsruiufyxhgxglebaebnnjndndy.org
peakconfor.org/20170116/index.html
alxqer.hk:80/k2jQPzNxSU/95fkOZExBQD.php?id=shakira@justiceto.com&num=465817589985325
pmenboeqhyrpvomq.yw4629.top
awpjugdubjlpwmpqsfjb.com
cdwagypboolcs.biz
linkd.in/1dwSnY1
stics.com/
www.ijeart.com/
news.academiccfp.com/submission.htm
loyabc.com:80/vfmc1mjskrtu/1apitn.php?id=kevin.baker@hbfutlers.com
w3.org/TR/html4/strict.dtd
oqwygprskqv65j72.1kh9ct.top
ueoiicsszefb.biz
yerconfole.org/20170115/unsubscribe.html
socotu.com.tn/
news.academiccfp.com/FPH.htm
tcsmith.com/
vyohacxzoue32vvk.0ayn1s.top
nie.edu.sg/
tkuceuah.net
laarmjndjkueaxfxbeewpptsxu.net
ciseng.org/2V4LLV/
socotu.com.tn/
clfctqge.net
chdiekgopartylines.com:80/XgnGV4/4YaUQyNdXA.php?id=mobile@fastestgeb.it
blog.interiextfilecasar.com:80/Z8iGogy/92oqcMLUphaeZv8.php?id=christopher.nolan@zoeaacs.com
schemas.microsoft.com/office/2004/12/omml
xpcx6erilkjced3j.16hwwh.top
ws.fasdistribuzione.com/WSRecuperaStatoSped/WSRecuperaStatoSped.asmx
pvjpfwlblergeex.net
aiche.org/community/awards/aiches-35-under-35-award
emailconnect.in/tl.php?p=fq5/d7m/rs/c7x/142/rs//http%3A%2F%2Fwww.samsung.com%2Fin%2Fconsumer%2Fmemory-storage%2Fssd%2F
h24info.o.kics.it/desabonnement?Key=0swedg3w9e00etn1uxfkODEzNzgxOTA1NzIxNDA3ODQx
wjtqjleommc4z46i.uwckha.top
qfjhpgbefuhenjp7.1225wj.top
hourrtxcouette.net:80/jW7hoALgfz/4qdVwfyGQe.php?id=koen.johen@bactla.lu
xyxzo.com:80/mhnqptq38x60/7fulexqqre.php?id=jimmy.page@nara.icaast.se

Problems


enems.blog.vitdogl.bg:80/l4Z1dqJLQ/3cfJ7iV.php?id=norah.jones@hotmail.com
loyabc.com:80/vfmc1mjskrtu/1apitn.php?id=kevin.baker@hbfutlers.com
h24info.o.kics.it/desabonnement?Key=0swedg3w9e00etn1uxfkODEzNzgxOTA1NzIxNDA3ODQx
mxweb.mnx.com/TrackOrder.aspx?TrackingNumber=6547794
monfredasas.com/administrator/components/com_acymailing/ex=
dokadfa.ru:80/mFovYD0C6pf/n1SH2k.php?id=peter.jackson@reynoraad.com
creiicreelectrique.net:80/akQfr5/z39otl1G.php?id=eric.johnson@belfasts.be
qfjhpgbefuhenjp7.1225wj.top
tcsmith.com/
snaduvhaphgxlawiwv.biz
wiki.apache.org/spamassassin/DnsBlocklists#dnsbl-block
laarmjndjkueaxfxbeewpptsxu.net
rsruiufyxhgxglebaebnnjndndy.org
w3.org/TR/html4/strict.dtd
peakconfor.org/20170116/index.html
clfctqge.net
nie.edu.sg/
vyohacxzoue32vvk.0ayn1s.top
fudlbiggughqfmxxieqlabsrriluecv.com
tkuceuah.net
socotu.com.tn/
pvjpfwlblergeex.net
schemas.microsoft.com/office/2004/12/omml
socotu.com.tn/
xpcx6erilkjced3j.16hwwh.top
www.ijeart.com/
unocl45trpuoefft.y721yz.top
aiche.org/community/awards/aiches-35-under-35-award
ueoiicsszefb.biz
news.academiccfp.com/FPH.htm
awpjugdubjlpwmpqsfjb.com
ciseng.org/2V4LLV/
news.academiccfp.com/submission.htm
cdwagypboolcs.biz
linkd.in/1dwSnY1
cftivusqaomzrwrfzgow.com
pmenboeqhyrpvomq.yw4629.top
stics.com/
yerconfole.org/20170115/unsubscribe.html
soxnkvfcqjjooef.org
ciseng.org/N60D22/
hourrtxcouette.net:80/jW7hoALgfz/4qdVwfyGQe.php?id=koen.johen@bactla.lu
chdiekgopartylines.com:80/XgnGV4/4YaUQyNdXA.php?id=mobile@fastestgeb.it
oqwygprskqv65j72.1kh9ct.top
news.academiccfp.com/journals.htm
wjtqjleommc4z46i.uwckha.top
news.academiccfp.com/10.5923.j.fph.20160606.03.htm
p27dokhpz2n7nvgr.12a63k.top
ws.fasdistribuzione.com/WSRecuperaStatoSped/WSRecuperaStatoSped.asmx
w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd
xyxzo.com:80/mhngptq38x60/7fulexqgre.php?id=jimmy.page@nara.icaast.se
blog.interiextfilecasar.com:80/Z8iGogy/92oqcMLUphaeZv8.php?id=christopher.nolan@zoeaacs.com
news.academiccfp.com/10.5923.j.cmaterials.20160606.03.htm

Problems


cruise.co.uk/news/?utm_campaign=NEWS220117&utm_medium=email&utm_source=NEWS220117&email=sh.smith@sdv.com
w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd
aiche.org/community/awards/aiches-35-under-35-award
qfjhpgbefuhenjp7.1225wj.top
wiki.apache.org/spamassassin/DnsBlocklists#dnsbl-block
snaduvhaphgxlawiwv.biz
news.academiccfp.com/10.5923.j.fph.20160606.03.htm
news.academiccfp.com/journals.htm
clfctqge.net
pvjpfwlblergeex.net
ws.fasdistribuzione.com/WSRecuperaStatoSped/WSRecuperaStatoSped.asmx
news.academiccfp.com/FPH.htm
loyabc.com:80/vfmc1mjskrtu/1apitn.php?id=kevin.baker@hbfutlers.com
ciseng.org/N60D22/
tcsmith.com/
yerconfole.org/20170115/unsubscribe.html
xyxzo.com:80/mhngptq38x60/7fulexqgre.php?id=jimmy.page@nara.icaast.se
supermissivefit.com/unsubscribe.php?M=1351510&C=54ae52d0497fadf40f48c0877bbdaa92&L=10&N=272
dokadfa.ru:80/mFovYD0C6pf/n1SH2k.php?id=peter.jackson@reynoraad.com
ozkaabfest.com:80/jP4pRS/4SqkMF2JDQsR.php?id=mc.hammer@registerliet.se
www.ijeart.com/
stics.com/
socotu.com.tn/
xpcx6erilkjced3j.16hwwh.top
ueoiicsszefb.biz
wjtqjleommc4z46i.uwckha.top
laarmjndjkueaxfxbeewpptsxu.net
ciseng.org/2V4LLV/
unocl45trpuoefft.y721yz.top
nie.edu.sg/
peakconfor.org/20170116/index.html
oqwygprskqv65j72.1kh9ct.top
rsruiufyxhgxglebaebnnjndndy.org
vyohacxzoue32vvk.0ayn1s.top
socotu.com.tn/
w3.org/TR/html4/strict.dtd
fudlbiggughqfmxxieqlabsrriluecv.com
news.academiccfp.com/submission.htm
monfredasas.com/administrator/components/com_acymailing/ex=
cdwagypboolcs.biz
mxweb.mnx.com/TrackOrder.aspx?TrackingNumber=6547794
pmenboeqhyrpvomq.yw4629.top
tkuceuah.net
p27dokhpz2n7nvgr.12a63k.top
schemas.microsoft.com/office/2004/12/omml
linkd.in/1dwSnY1
news.academiccfp.com/10.5923.j.cmaterials.20160606.03.htm
click.e.vineyardvines.com/?qs=6ba2b95d7aec369a2e70c9a04d61348a7684b633ffe55dc996a089b89c628b325075780cc9268cf4
soxnkvfcqjjooef.org
h24info.o.kics.it/desabonnement?Key=0swedg3w9e00etn1uxfkODEzNzgxOTA1NzIxNDA3ODQx
navi.mail.carenet.com/c0/m1/teid-t9YQTYIV8TB62mDtu2Owc7LmIboTQU1UdeSSV4/www.carenet.com/news/general/carenet/43305
cftivusqaomzrwrfzgow.com
awpjugdubjlpwmpqsfjb.com

Problems


wiki.apache.org/spamassassin/DnsBlocklists#dnsbl-block
mxweb.mnx.com/TrackOrder.aspx?TrackingNumber=6547794
w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd
aiche.org/community/awards/aiches-35-under-35-award
hourrtxcouette.net:80/jW7hoALgfz/4qdVwfyGQe.php?id=koen.johen@bactla.lu
h24info.o.kics.it/desabonnement?Key=0swedg3w9e00etn1uxfkODEzNzgxOTA1NzIxNDA3ODQx
monfredasas.com/administrator/components/com_acymailing/ex=
creiicreelectrique.net:80/akQfr5/z39otl1G.php?id=eric.johnson@belfasts.be
ws.fasdistribuzione.com/WSRecuperaStatoSped/WSRecuperaStatoSped.asmx
laarmjndjkueaxfxbeewpptsxu.net
ueoiicsszefb.biz
fudlbiggughqfmxxieqlabsrriluecv.com
news.academiccfp.com/submission.htm
qfjhpgbefuhenjp7.1225wj.top
ciseng.org/N60D22/
news.academiccfp.com/journals.htm
snaduvhaphgxlawiwv.biz
cftivusqaomzrwrfzgow.com
vyohacxzoue32vvk.0ayn1s.top
unocl45trpuoefft.y721yz.top
socotu.com.tn/
rsruiufyxhgxglebaebnnjndndy.org
wjtqjleommc4z46i.uwckha.top
stics.com/
nie.edu.sg/
ciseng.org/2V4LLV/
pmenboeqhyrpvomq.yw4629.top
www.ijeart.com/
yerconfole.org/20170115/unsubscribe.html
xpcx6erilkjced3j.16hwwh.top
schemas.microsoft.com/office/2004/12/omml
tcsmith.com/
pvjpfwlblergeex.net
soxnkvfcqjjooef.org
news.academiccfp.com/10.5923.j.fph.20160606.03.htm
p27dokhpz2n7nvgr.12a63k.top
peakconfor.org/20170116/index.html
linkd.in/1dwSnY1
w3.org/TR/html4/strict.dtd
oqwygprskqv65j72.1kh9ct.top
awpjugdubjlpwmpqsfjb.com
news.academiccfp.com/FPH.htm
loyabc.com:80/vfmc1mjskrtu/1apitn.php?id=kevin.baker@hbfutlers.com
clfctqge.net
tkuceuah.net
enems.blog.vitdogl.bg:80/l4Z1dqJLQ/3cfJ7iV.php?id=norah.jones@hotmail.com
ozkaabfest.com:80/jP4pRS/4SqkMF2JDQsR.php?id=mc.hammer@registerliet.se
supermissivefit.com/unsubscribe.php?M=1351510&C=54ae52d0497fadf40f48c0877bbdaa92&L=10&N=272
cdwagypboolcs.biz
alxqer.hk:80/k2jQPzNxSU/95fkOZExBQD.php?id=shakira@justiceto.com&num=465817589985325
socotu.com.tn/
xyxzo.com:80/mhngptq38x60/7fulexqgre.php?id=jimmy.page@nara.icaast.se

Machine Learning

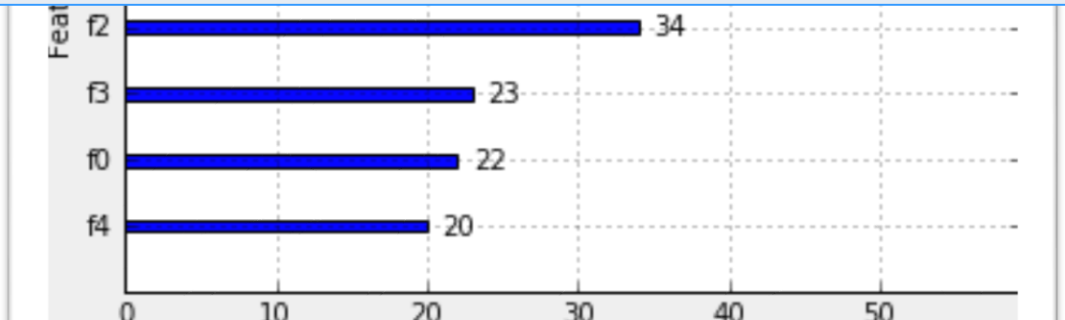
XGBoost, SVM, Random Forest, ...

```
trainx, testx, trainy, testy = train_test_split(X, Y, test_size=0.3, random_state=7)
trainset, testset = xgb.DMatrix(trainx, trainy), xgb.DMatrix(testx)

params = {'max_depth':10, 'eta':1, 'silent':1, 'objective':'binary:logistic' }
model = xgb.train(params, trainset, num_boost_round=10)

predictions = model.predict(testset)
```

```
booster[0]:
0:[f1<127.5] yes=1,no=2,missing=1
        1:[f7<28.5] yes=3,no=4,missing=3
                3:[f5<30.95] yes=7,no=8,missing=7
                        7:[f0<5.5] yes=15,no=16,missing=15
                                15:leaf=-1.89091
                                16:leaf=-0.5
                        8:[f6<0.9045] yes=17,no=18,missing=17
```



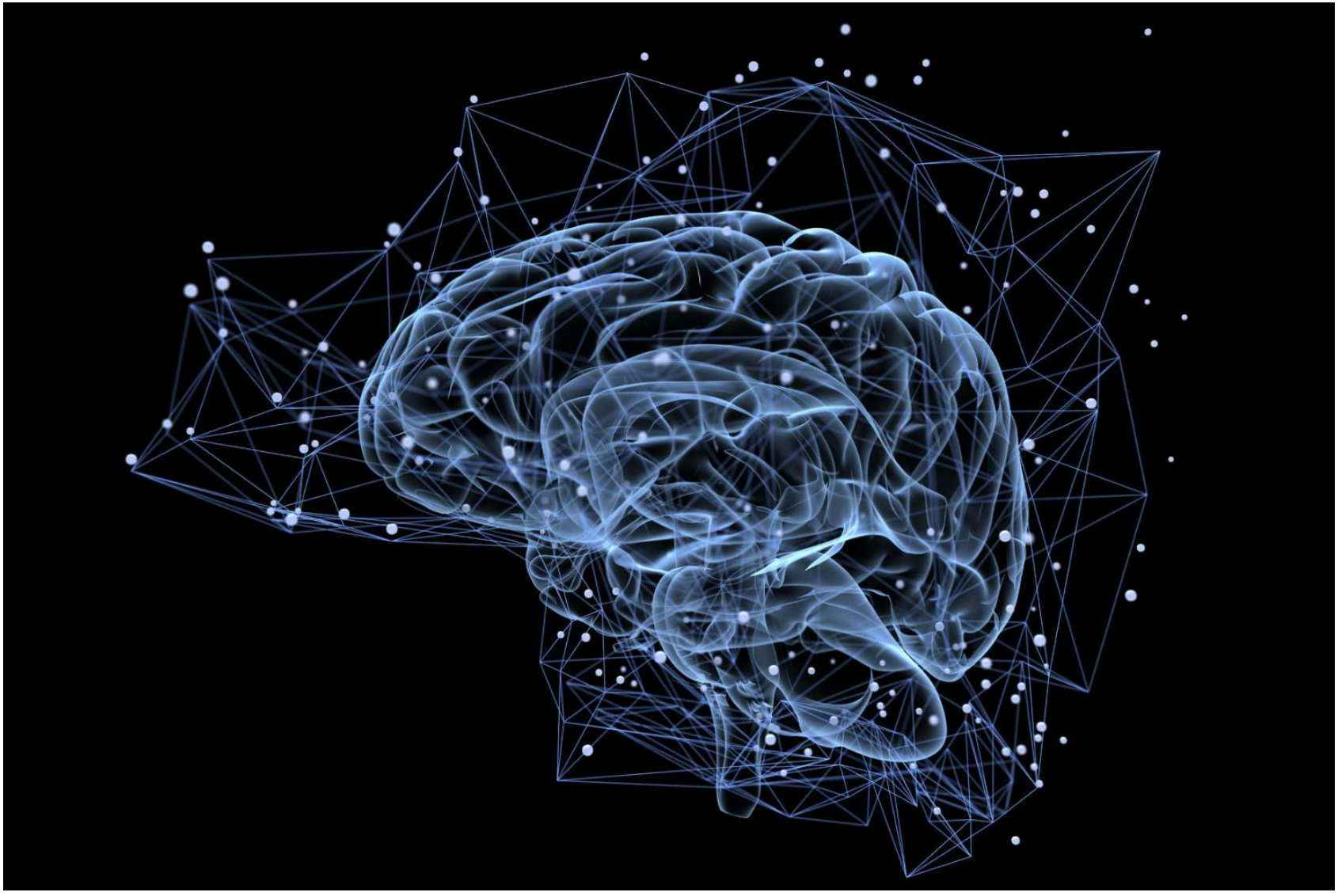
Would Traditional ML Work?

XGBoost/SVM/RandomForest

- Pros
  - Easy to program
  - Very fast training
  - Perform well against tabular input data
  - Use human intelligence and heuristics
- Cons
  - Hard to debug when it mis-predicts
  - Development cost is high
  - Feature engineering is required

Deep Learning

Neural Network



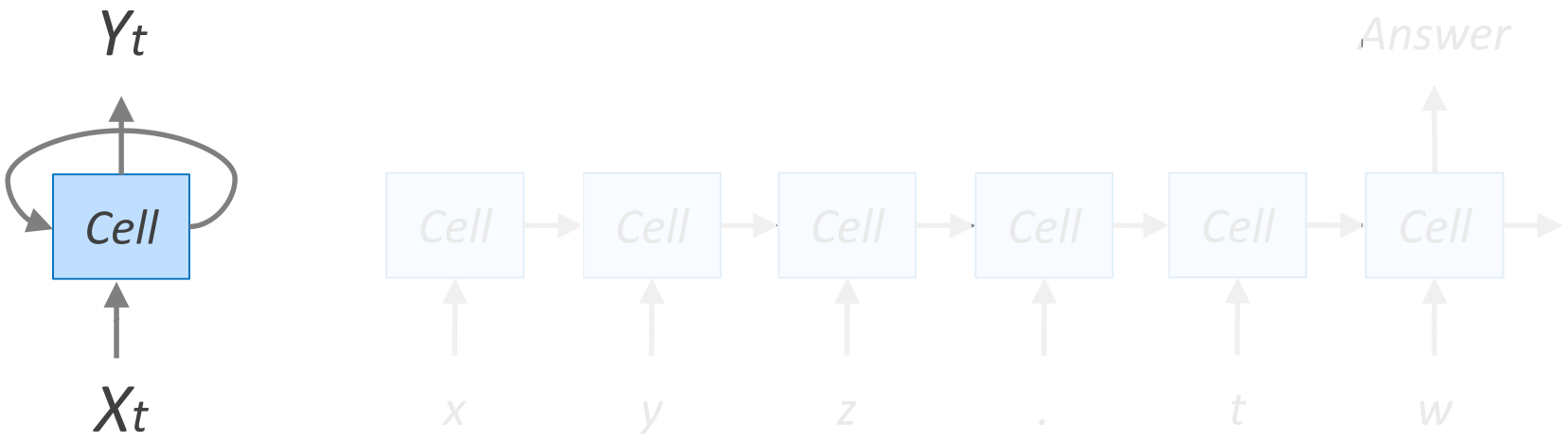
Malicious URL Detection

```
Tensorboard logs dir for this run is model/checkpoint/run1
Model loaded in 2.6 sec
Classification finished in 5.2 sec
Result saved in 0.0 sec
mail.loft.com/a/tBYhQ9GBgYd1XB9XjnHNuJ8T0ct/freeship?EMAIL_HASH=c0f53badcf51dd5ae65050f2d95b6f1c&email=sudheer.kumar@infogain.com legitimate
view.email.vegas.com/?qs=44f8e735cc62ba1625f82687634b3ae63d67ec1e6f520b39cb1988190a4f55eba48987b421e771d4023bb17fb95994de legitimate
www.facebook.com/n/?Stephwalker1%2Fposts%2F10158005332385548&comment_id=10158011284830548&aref=1485117733209677&medium=email&mid=546b4a15a03
ode=1.1485117732.AbnpAWrxdSMDLMD2&n_m=kelly.sprague%40cengage.com legitimate
tigerdirect.com/email/3WWEBEML653.asp?MobileOptOut=1&SRCCODE=3WWEBEML653&utm_source=EML&utm_medium=main&utm_campaign=3WWEBEML653&elqTrackId=
5&elq=8a5233a641db4b82ab318b2dbd74d514&elqaid=628&elqat=1&elqCampaignId=489 legitimate
1onabcf.ru:80/yPY2iIcC/3oVS15.php?id=peter.jackson@pfipaer.com cryptolocker
rlqjhwmu.com locky
icjwdktucoaio.com locky
click.email.vegas.com/?qs=16bc3fc881c351af16ea2d946[illegible]6802ff4b1c72 legitimate
www.facebook.com/ legitimate
p27dokhpz2n7nvgr.1cglxz.top cerber-rigek
p27dokhpz2n7nvgr.12smak.top cerber-rigek
w3.org/TR/html4/loose.dtd legitimate
www.tahitinuitravel.biz/ legitimate
www.rivieraholidays.in/ legitimate
www.tahiti-infos.com/Intemperies-a-Tahiti-la-piste-aeroportuaire-est-fermee-jusqu-a-nouvel-ordre_a157036.html legitimate
www.w3.org/TR/xhtml1/DTD/xhtml1-transitional.dtd legitimate
www.spiceworks.com/ legitimate
www.facebook.com/n/?kjohannesdottir1%2Fposts%2F10211634755642816&comment_id=10211635744347533&aref=1485111640385846&medium=email&mid=546b3b4
7&bcode=1.1485117353.AbnjnjRtQytXnFOy&n_m=margret%40valitor.is&lloc=1st_cta legitimate
www.vinnumalastofnun.is/ legitimate
4kqd3hmqgptupi3p.wz139z.top cerber-rigek
27lelchgcvs2wpm7.b7mciu.top cerber-rigek
ffoqr3ug7m726zou.do9wwg.top cerber-rigek
www.husbot.is/ legitimate
www.boksala.is/valitorpaymentgateway/payment/response/?Kortategund=VISA&KortnumerSidustu=7172&KortnumerStjarnad=415552******7172&Dagsetning=
9&Faerslunumer=0384273&VefverslunSalaID=c3a61021-9998-4e06-bc91-d52b74c87cb2&Tilvisunarnumer=200004785&RafraenUndirskriftSvar=0aa4a3e88b3517
```

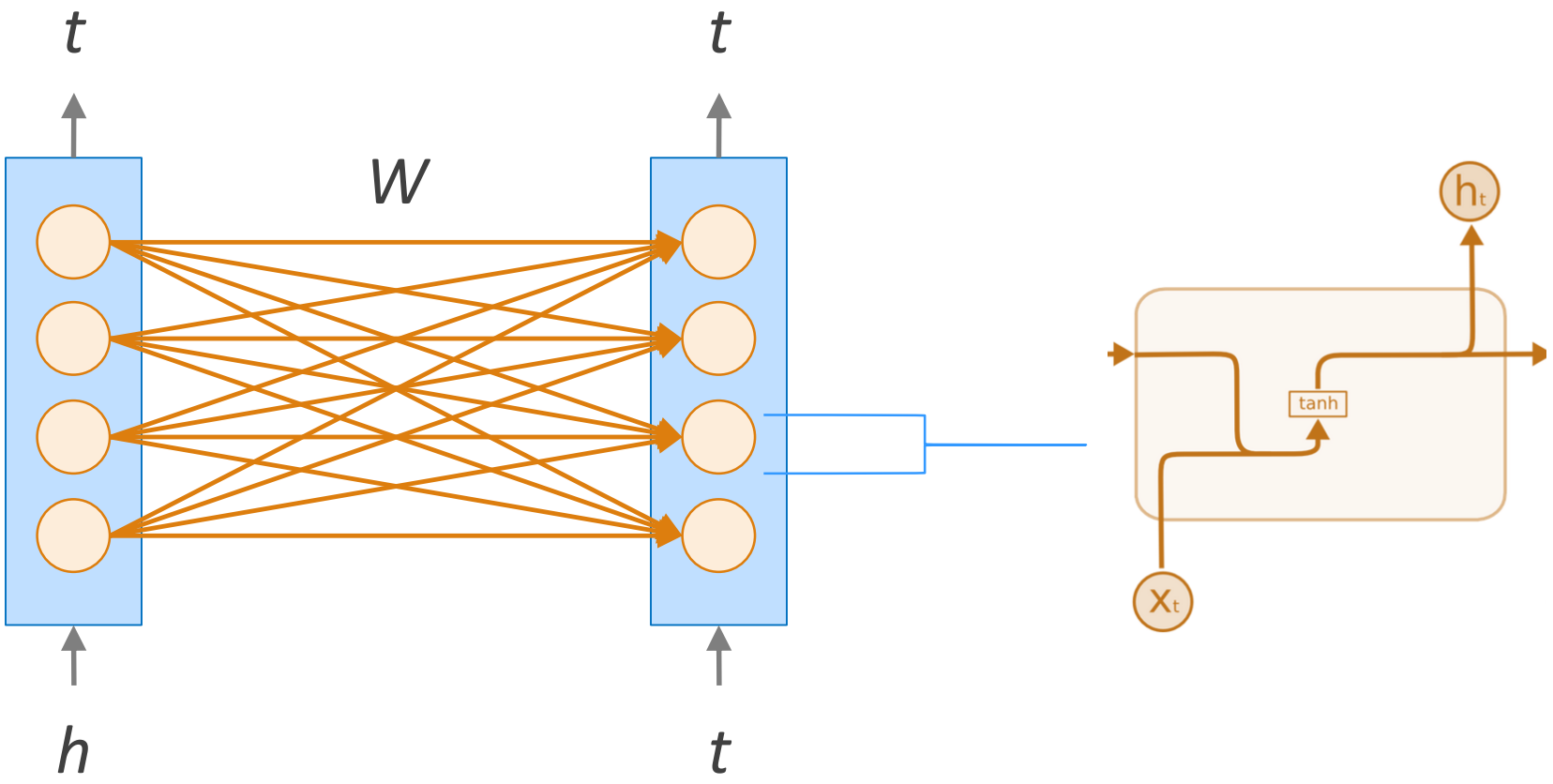
DEMO

Deep Learning

Recurrent neural network (RNN)

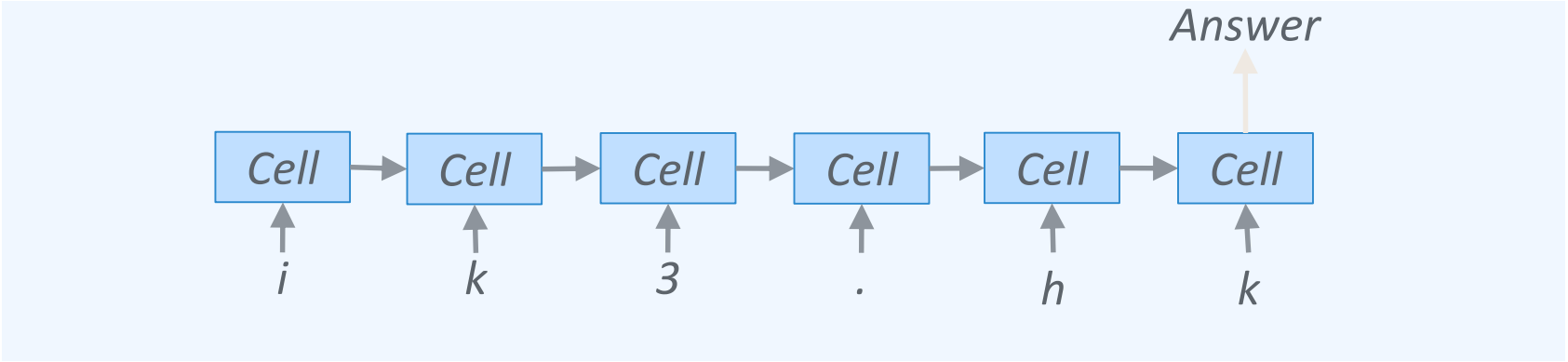


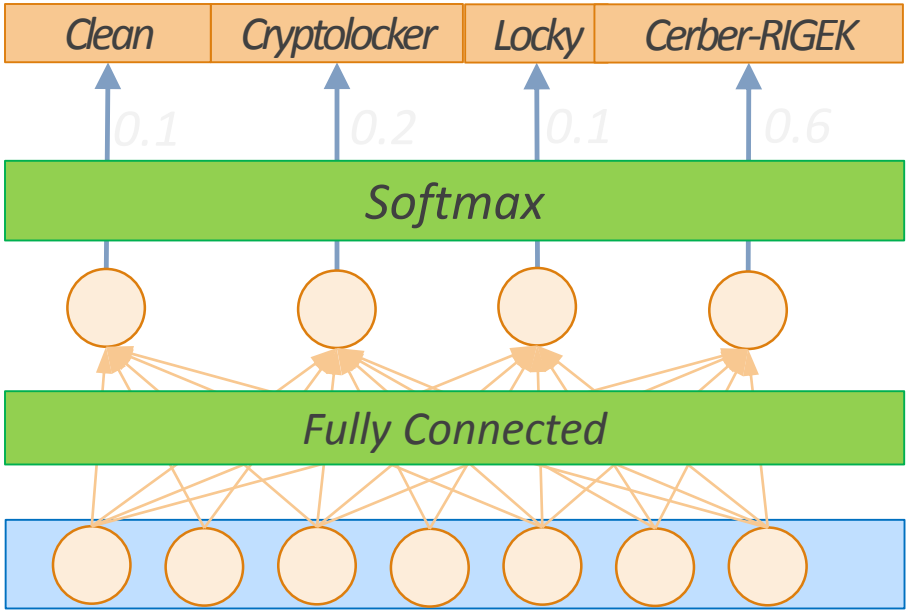
RNN Cell and Connections



Classification Using RNN Cell

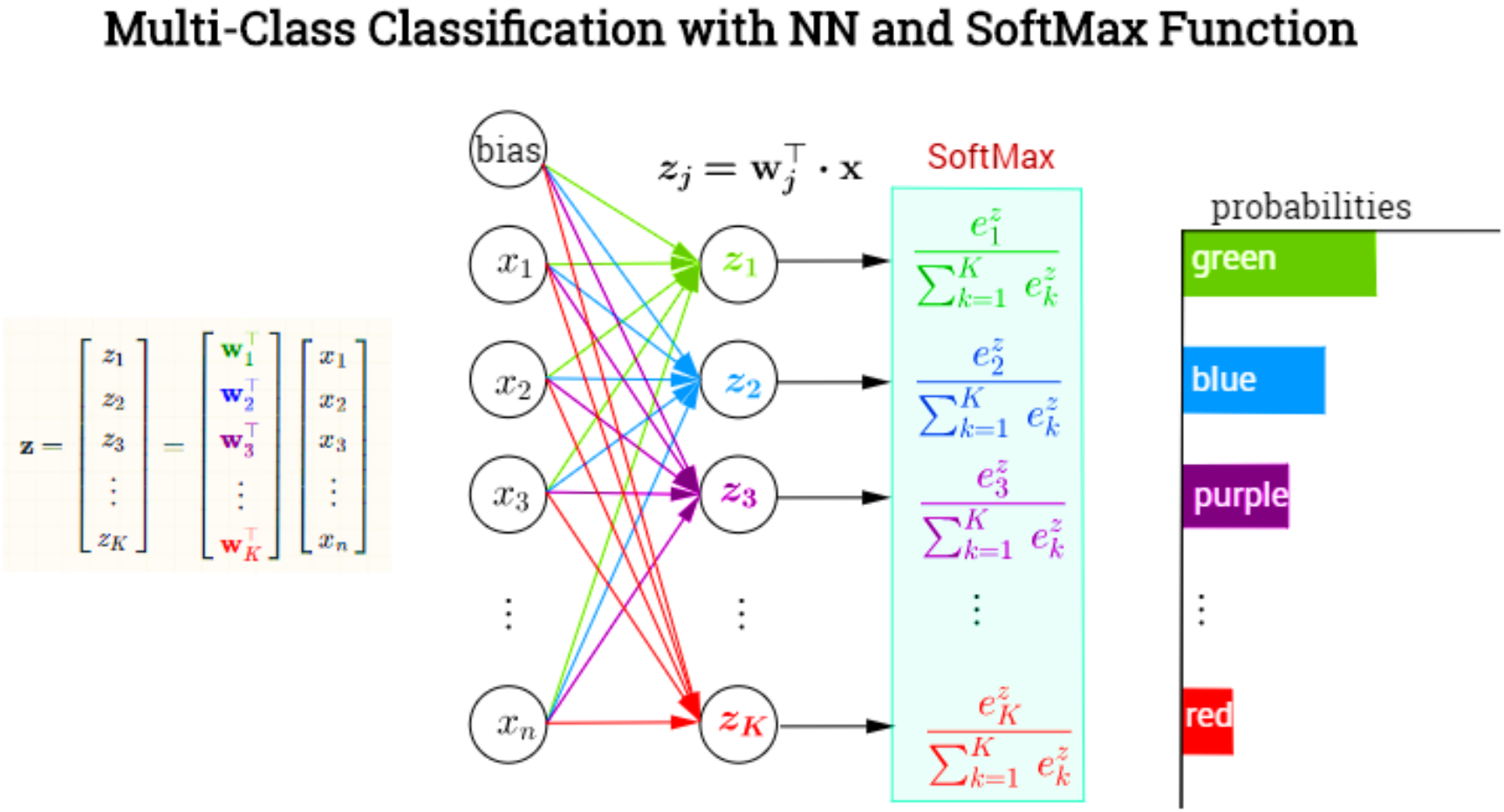
Cost: softmax cross entropy





Classification Using RNN Cell

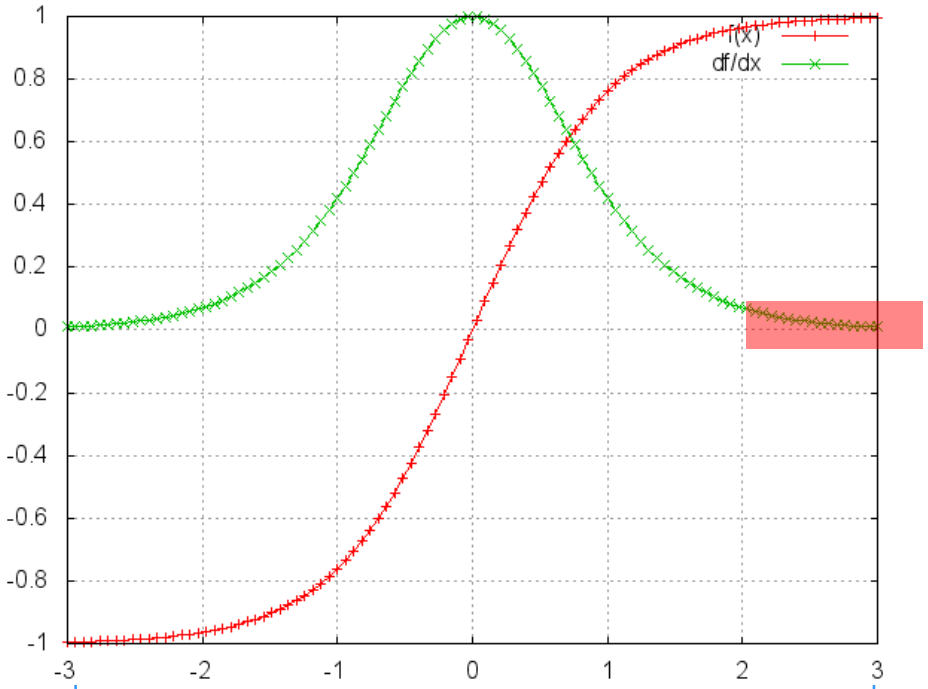
Training: Back Propagation Through Time



Source: https://stats.stackexchange.com/questions/273465/neural-network-softmax-activation

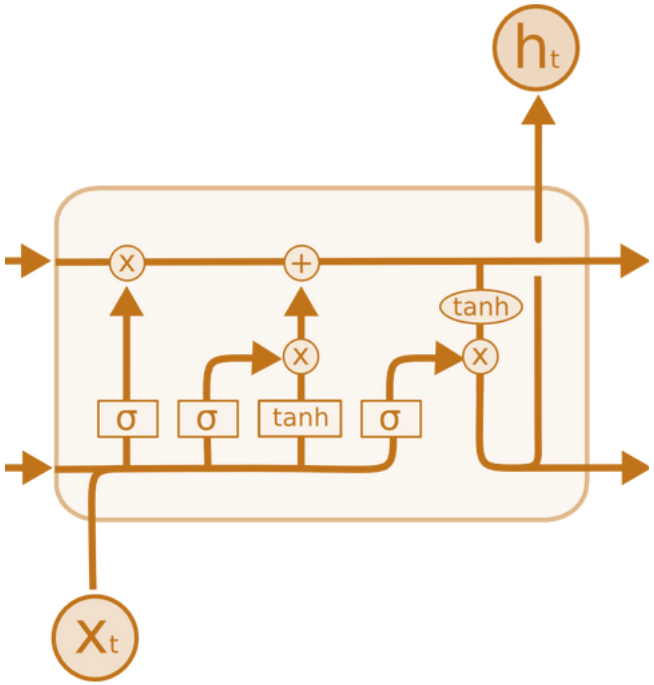
RNN Cell

Vanishing Gradient Problem



LSTM Cell

Long Short Term Memory



$$X = X_t \mid H_{t-1}$$

$$f = \sigma(X.W_f + b_f)$$

$$u = \sigma(X.W_u + b_u)$$

$$r = \sigma(X.W_r + b_r)$$

$$X' = \tanh(X.W_c + b_c)$$

$$C_t = f * C_{t-1} + u * X'$$

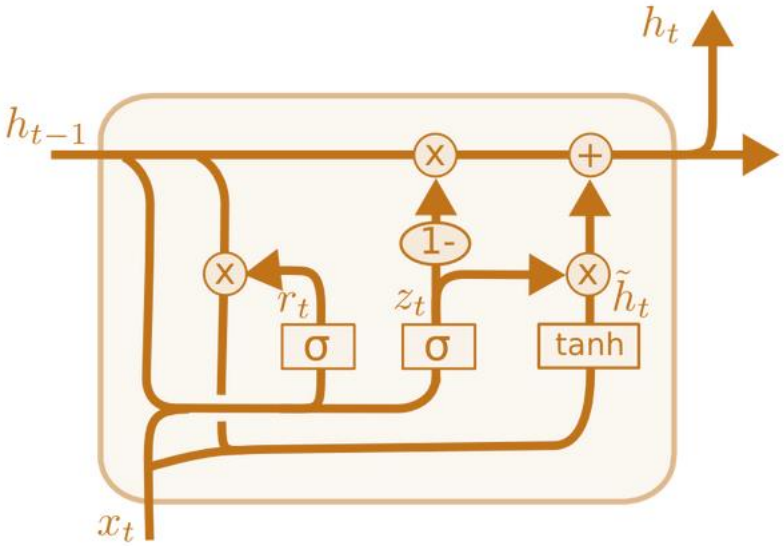
$$H_t = r * \tanh(C_t)$$

$$Y_t = \text{softmax}(H_t.W + b)$$

Source: http://colah.github.io/posts/2015-08-Understanding-LSTMs/

GRU Cell

Gated Recurrent Unit



$$z_t = \sigma\left(W_z \cdot [h_{t-1}, x_t]\right)$$

$$r_t = \sigma\left(W_r \cdot [h_{t-1}, x_t]\right)$$

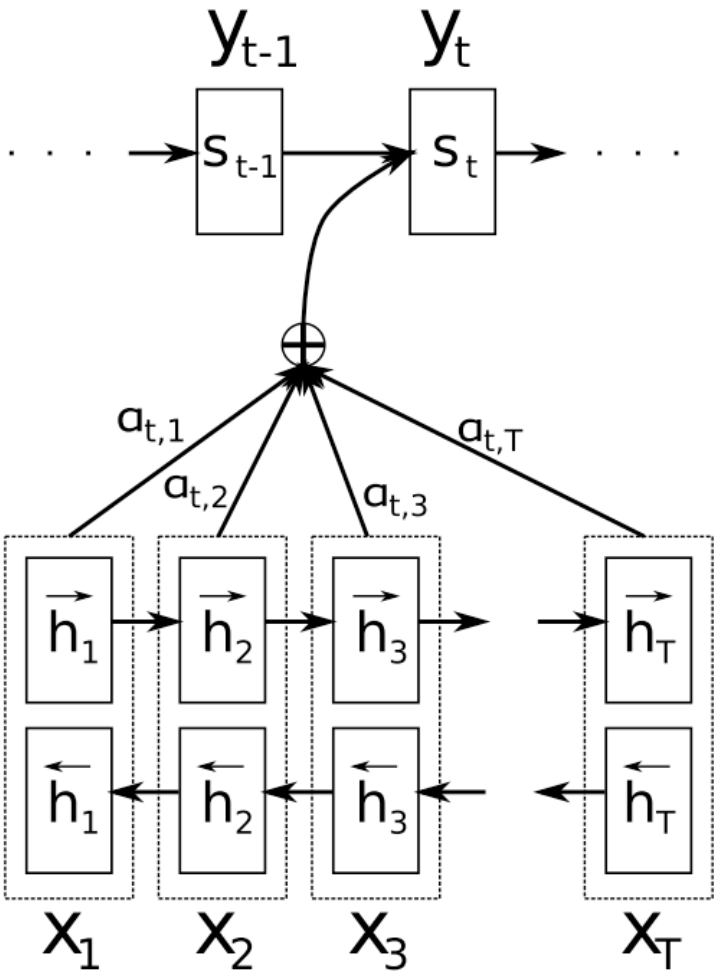
$$\tilde{h}_t = \tanh\left(W \cdot [r_t * h_{t-1}, x_t]\right)$$

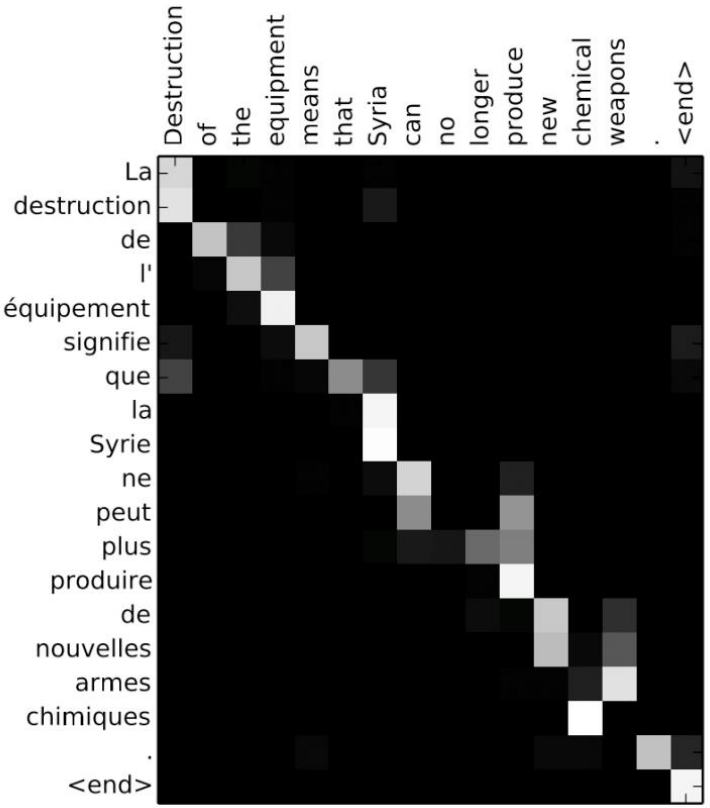
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Source: http://colah.github.io/posts/2015-08-Understanding-LSTMs/

RNN with Attention

Example: Neural Machine Translation (NMT)





Source: https://arxiv.org/pdf/1409.0473.pdf

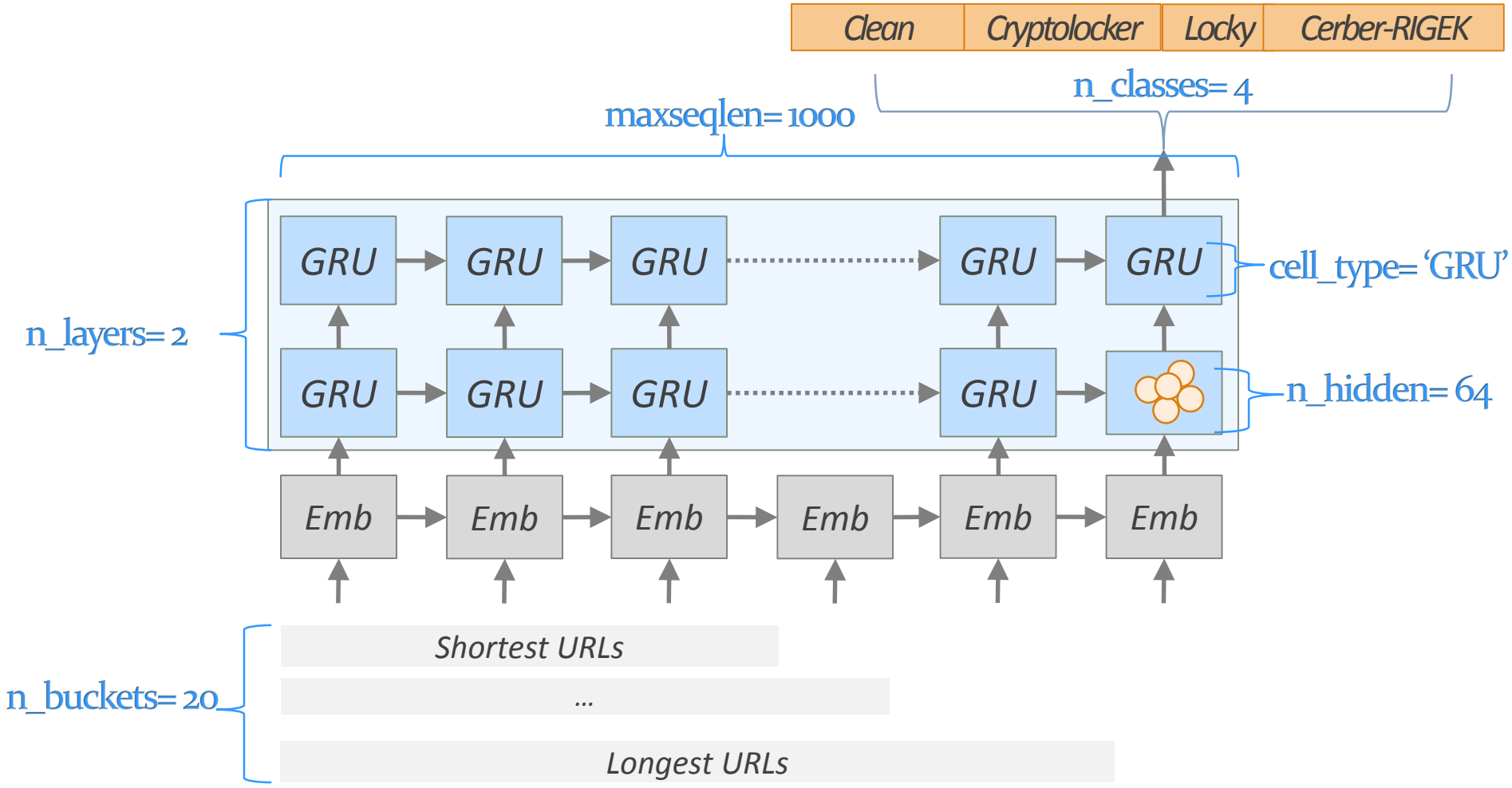
Embedding

- Symbols do not carry their natural semantics within them whereas continuous signals such as audios and videos do.

www.facebook.com/n/?kjoha

Neural Blacklist Network

Architecture & Hyper parameters



Feature & Dataset

- Feature

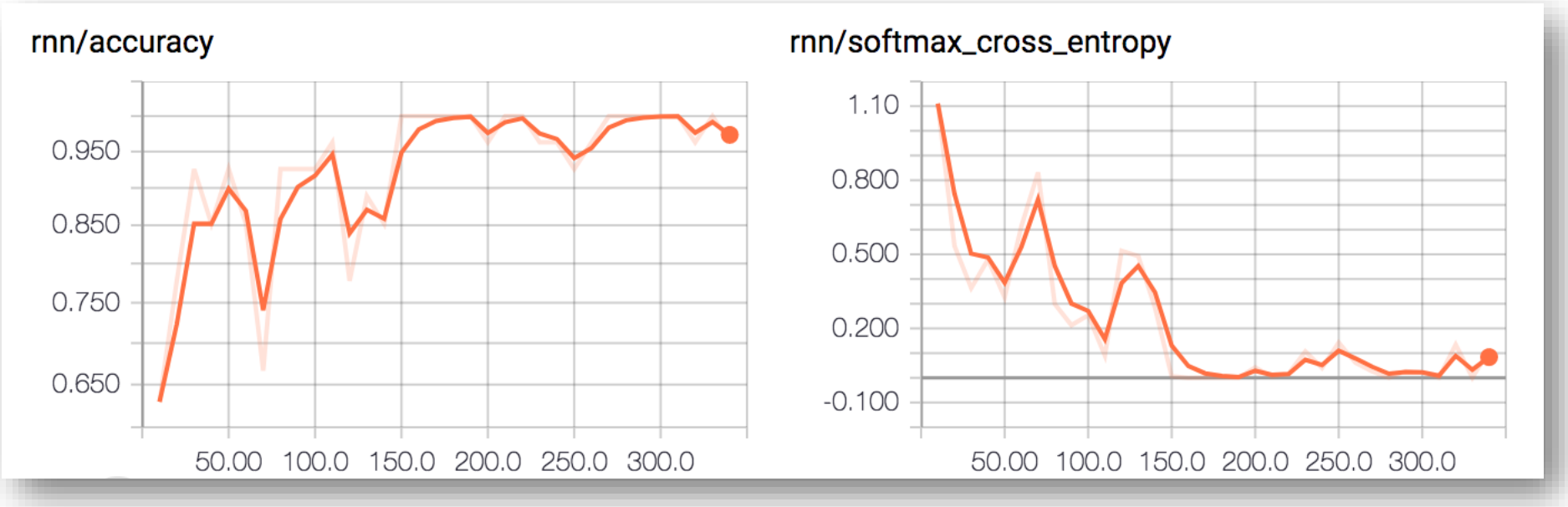
```
>> batch_x[2]
array([101,  46, 119, 101, 116, 115, 101,  97, 108, 110, 101, 119, 115,
       108, 101, 116, 116, 101, 114,  46,  99, 111, 109,  47, 113,  47,
        74,  83,  56, 104,  87,  74, 108,  69, 111, 107, 119, 100,  69,
        73, 106, 115,  99,  81, 109,  88, 116, 102,  95, 115,  66,  48,
       122,  69,  72,  83, 119,  99, 110,  52,  55, 104, 105,  86,  97,
        55,  87,  74,  45,  76,  49,  74,  56,  81, 113, 112, 122, 118,
       105, 117, 106,  69,  86,   0,   0,   0,   0,   0,   0,   0,   0], dtype=int32)
>> tochar(batch_x[2])
'e.wetsealnewsletter.com/q/JS8hWJlEokwdEIjscQmXtf_sB0zEHSwcn47hiVa7WJ-L1J8QqpzviujEV\x00\
```

- Dataset
  - Sourcing
    - Legitimate URLs: Akamai log
    - Cryptolocker: Malware operations team
    - Locky v2/ Cerber-RIGEK : Ransomware tracker
  - Splits
    - train : validation : test = 0.1 : 0.1 : 0.8

RNN Model Space Analysis

Tensorboard

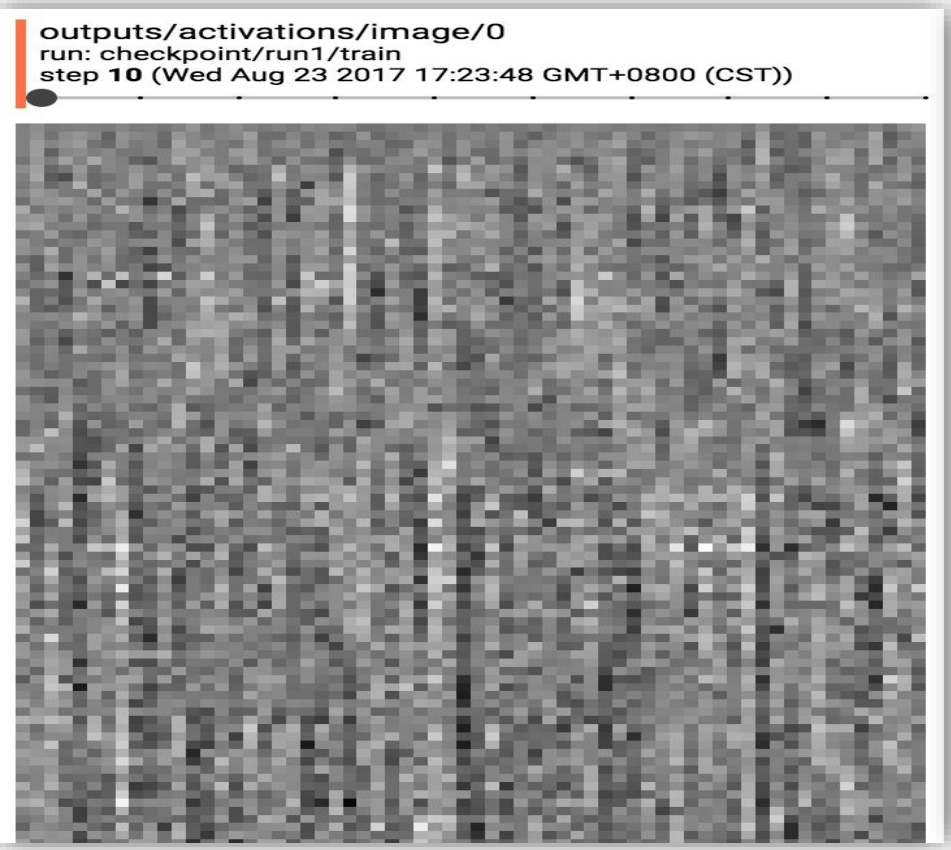
Accuracy & Cost



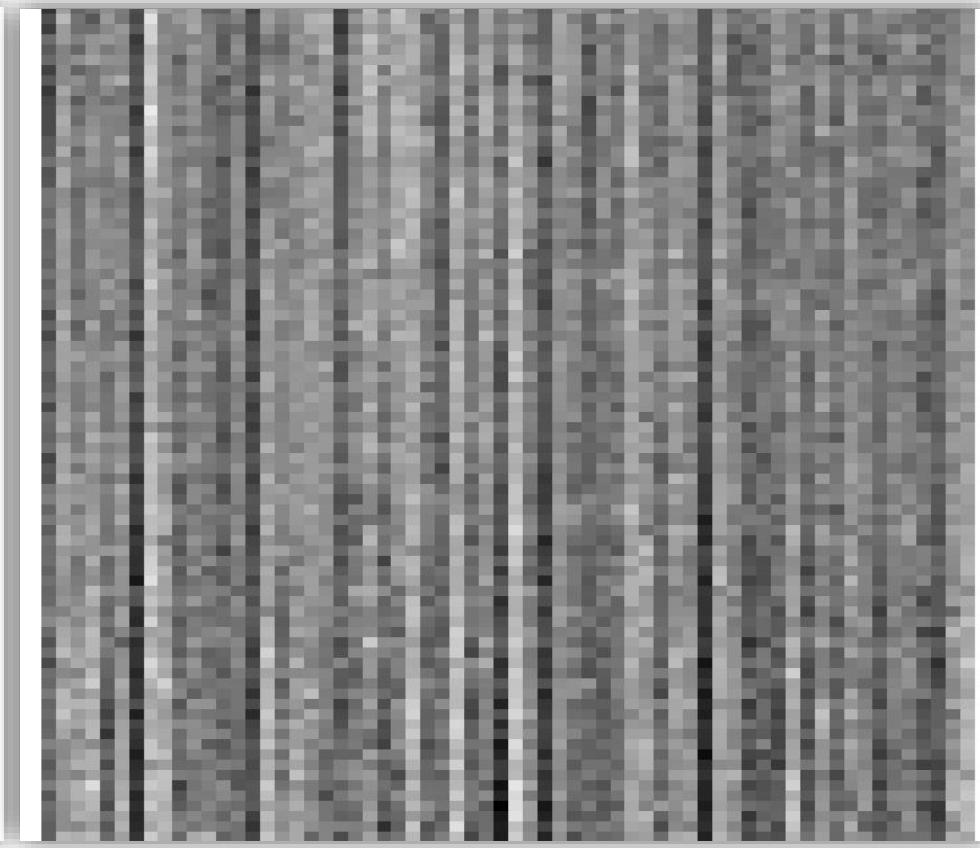
Tensorboard

Network States

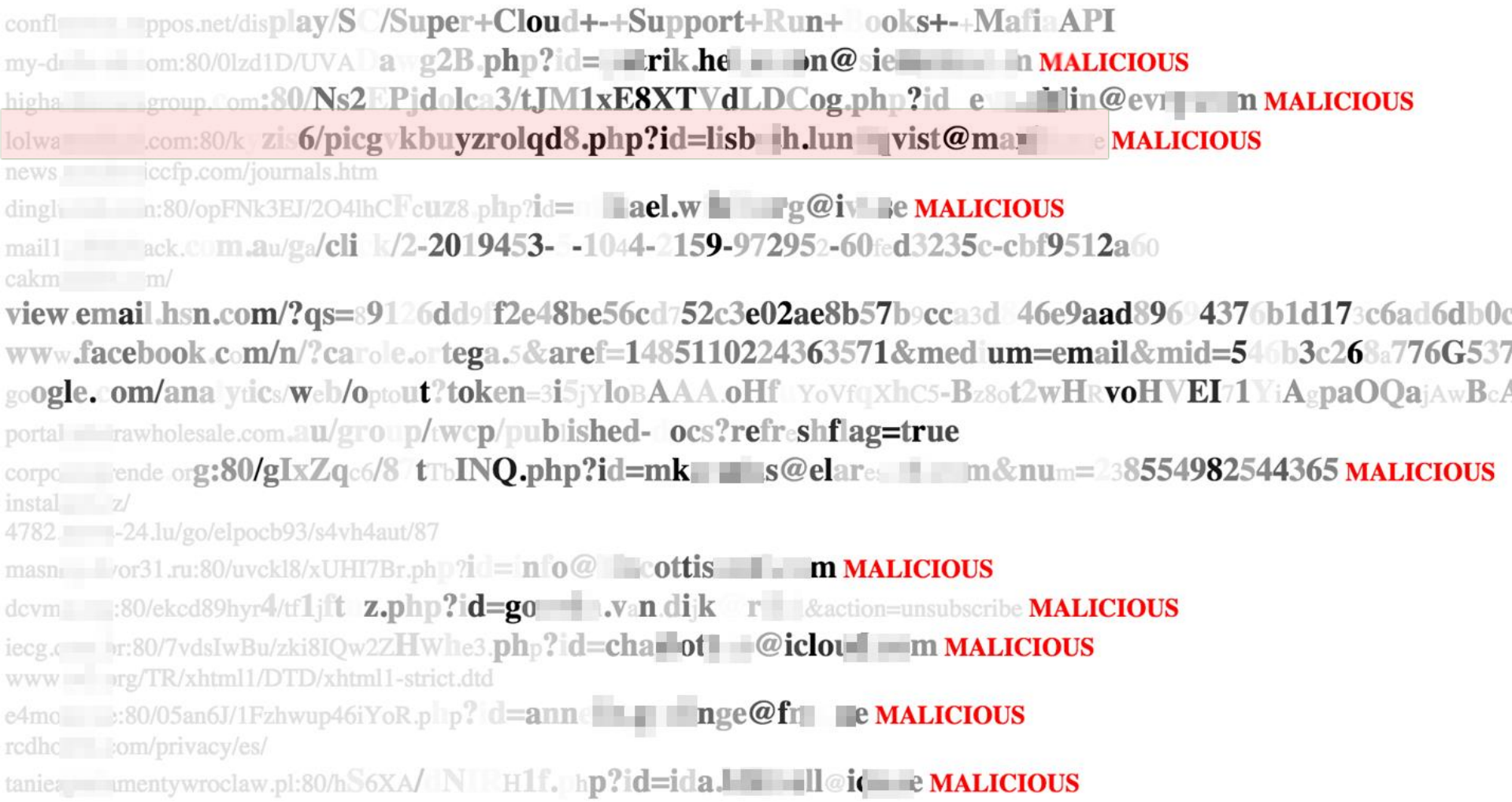
Before Training



After Training



CryptoLocker URL detection using attention



Experiment Summary

```
epoch 102 iteration 2492: cost 0.000037 (minibatch accuracy 100.000000%) [0 2 3]=[16  9  2]
epoch 102 iteration 2493: cost 0.000643 (minibatch accuracy 100.000000%) [0 2]=[ 4 23]
epoch 102 iteration 2494: cost 0.001896 (minibatch accuracy 100.000000%) [0 2]=[ 1 26]
epoch 102 iteration 2495: cost 0.000000 (minibatch accuracy 100.000000%) [0 1]=[ 3 24]
epoch 102 iteration 2496: cost 0.000000 (minibatch accuracy 100.000000%) [0 1]=[ 4 23]
epoch 102 iteration 2497: cost 0.000008 (minibatch accuracy 100.000000%) [0 2]=[ 6 21]
epoch 102 iteration 2498: cost 0.000000 (minibatch accuracy 100.000000%) [0]=[27]
epoch 102 step 99: validation accuracy 100.000000%
epoch 102 iteration 2499: cost 0.000000 (minibatch accuracy 100.000000%) [0]=[27]
epoch 102 iteration 2500: cost 0.000000 (minibatch accuracy 100.000000%) [0 1]=[18  9]
Training finished in 2813.9s
Best validation accuracy 100.000000%
Optimization complete with best validation accuracy 100.000000%
Training finished 2500 iterations in 2813.90 sec
validation accuracy 100.000000%
last_hidden_state.train: shape (2700, 64)
prediction: shape (2700, 4)
embedding metadata path: last_hidden_state.train.tsv
last_hidden_state.validation: shape (2700, 64)
prediction: shape (2700, 4)
embedding metadata path: last_hidden_state.validation.tsv
saving checkpoint iteration 0
```

Is it perfect?

- Undetected URL from test-cryptolocker.txt

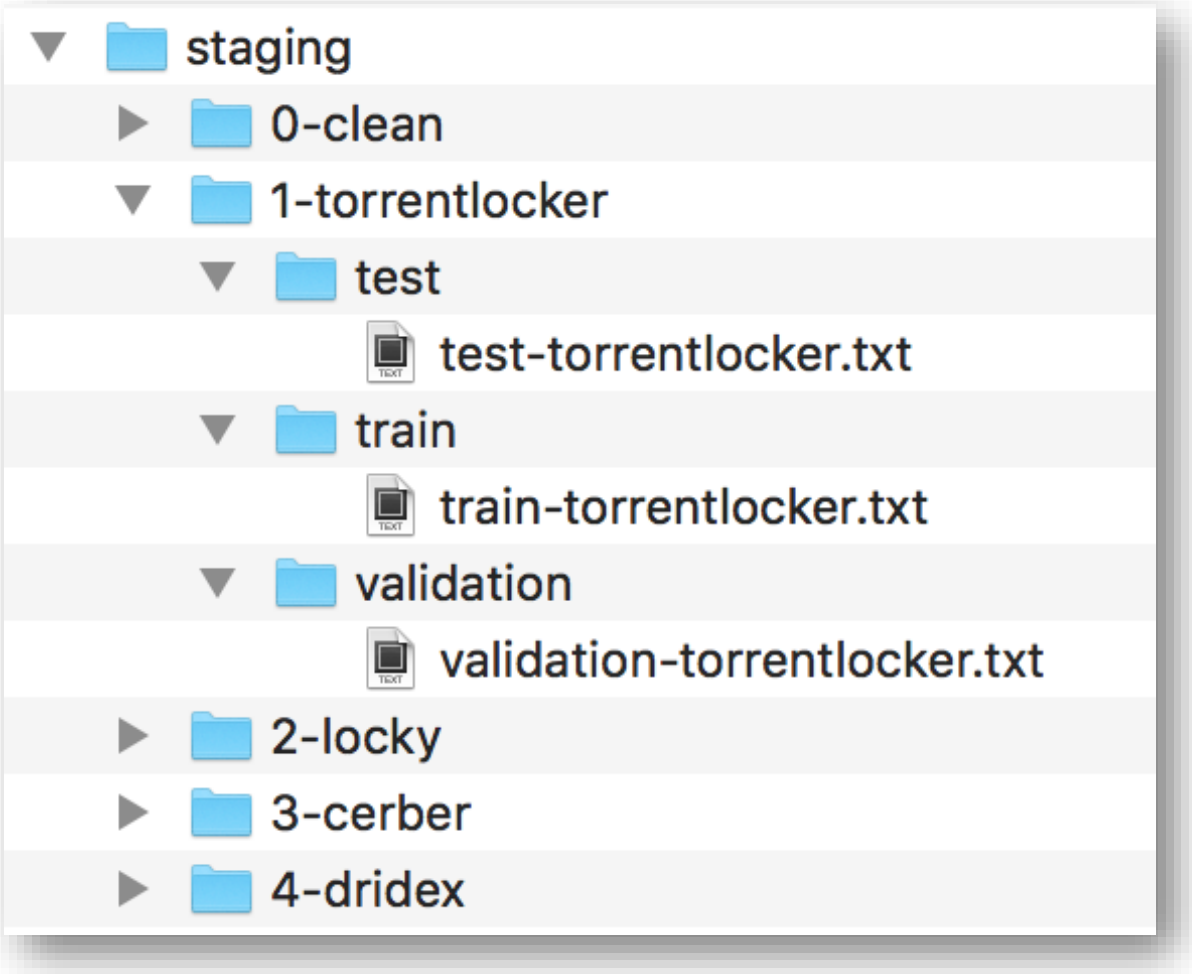
```
www.leriov.com:80/leriov3/player1.php?id=aH!BeF0cHM6
Ly9waG90b3MuZ29vZ2xlLmNvbS9zaGFyZS9B!BeFjF!BeFaXBNOH
B!BeFcmplbEEydU!BeFPX3ZZQTBLel!BeFKdjNmWVItMUFaM1UxQ
1UtX25oWDho!BeFjNTaDh!BeFaEs0bF85WXNlYVVySUNBP2tleT1
NMDV3YjB!BeFelNtVXpj@bfgo2TFdKVk1YQX!BeFjWHBOY0VKdGF
FdElhMHBS&id2=
```

- Analysis

This URL was misplaced in the test cryptolocker sample list. So this missed detection is a correct behaviour.

Towards Production

- Training and Testing by the samples



References

- http://www.wildml.com/2016/01/attention-and-memory-in-deep-learning-and-nlp/
- https://ransomwaretracker.abuse.ch/
- https://arxiv.org/pdf/1412.7449v3.pdf
- https://arxiv.org/abs/1409.0473
- https://magenta.tensorflow.org/2016/07/15/lookback-rnn-attention-rnn
- https://theneuralperspective.com/2016/11/20/recurrent-neural-network-rnn-part-4-attentional-interfaces/
- https://github.com/tensorflow/tensorflow/blob/master/tensorflow/contrib/rnn/python/ops/rnn_cell.py
- https://github.com/tensorflow/tensorflow/issues/4427
- http://r2rt.com/recurrent-neural-networks-in-tensorflow-iii-variable-length-sequences.html