

# Phishing Web Sites Features Classification Based on Extreme Learning Machine

<sup>1</sup>Naga Karthik Kosuri, <sup>2</sup>Nagasri B

<sup>1</sup>UG Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai.

<sup>2</sup>Assitant Professor, Department of Information and Technology, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai.

## Article Info

Volume 83

Page Number: 1222 - 1225

Publication Issue:

March - April 2020

## Abstract

Phishing is one of cybercrime's most widely perceived and risky ambushes. The aim of these attacks is to take the personal information and relationships used to organize trades. Phishing localities contain different signs within their information based on substance and web system. Extreme Learning Machine (ELM) based set for 30 features fusing Phishing Websites Data into UC Irvine Machine Learning Repository database is the reason behind this review. For the assessment of performance, ELM and other AI techniques, such as Support Vector Machine (SVM), Naïve Bayes (NB), were distinguished and considered to have the highest accuracy.

**Keywords:** *Extreme Learning Machine, Cyber Crimes, Support Vector Machine, Features Classification, Information Security, Phishing*

## Article History

Article Received: 24 July 2019

Revised: 12 September 2019

Accepted: 15 February 2020

Publication: 14 March 2020

## 1. Introduction

Web Usage has become a must bit of our step by step practices in light of rapidly creating advancement. In light of this snappy improvement of development and heightened usage of cutting edge structures, these devices have Data Security expanded phenomenal hugeness. The basic goal of maintaining security in advancement of information is to ensure that substantial protections are taken against threats and danger to be looked at by consumers through use of these developments.

Phishing is described as imitating strong locations with the aim of getting prohibitive information to destinations for various purposes every day, Usernames, passwords and nationality numbers for example. Phishing destinations contain various signs inside their information based on content and web software. Individuals providing the intimidation send the false site or email details to the target area just as it starts from an acquaintance, bank or whatever other reliable source that carries out strong trades. Substance of the website or email fuse asking

people to enter or reinforce their own details or change their passwords similarly as associations with locales that appear just as exact of the destinations of the affiliations concerned.

## 2. Existing System

The basic goal of maintaining security in propellers of knowledge is to ensure that critical steps of well-being are taken against risks and dangers that consumers will be exposed to when using these advances. Phishing is described as imitating trustworthy locations for accessing the prohibitive details that went to destinations every day for various purposes, such as usernames, passwords and numbers of citizenship. Phishing locales contain various bits of knowledge among their substance and web program based information. Individuals providing the blackmail will give the fake site or email details to the target area just as it starts from an affiliate, bank, or whatever other reliable source conducting strong trades. Substance of the web or email adds significance to sales in order to attract individuals to join or relive their own

details or to change their passwords similarly to connections with destinations that are just as true to the positions of the relevant affiliations.

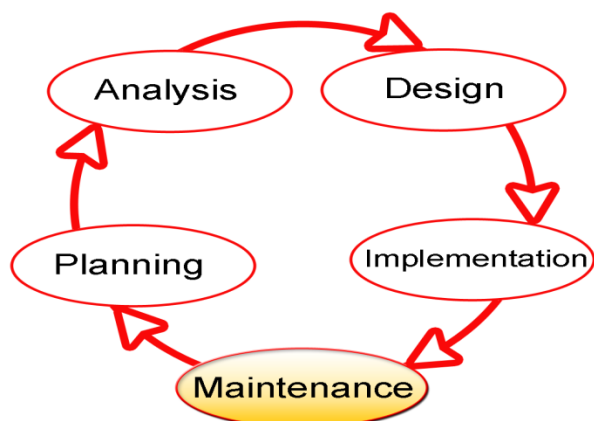
### 3. Proposed System

Along these lines study the issue of anticipating online purchase transformations in an internet business website. To comprehend user behavior and expectation on the web, existing indicators influence the conventional hunt example of entering queries then tapping on intriguing outcomes. In any case, transformation takes in excess of a tick. That is, after more than once clicking around and being presented to publicizing (i.e., retargeted), clients' ultimate success metric of the commercial center hunt is purchasing items. Past the customary instrument, our commitment is to permit the indicators to consider dynamic commercial center components for a more profound expectation of the two ticks and buys. In particular, motivated by customary pursuit issues we center around two research questions: "Expectation from market" and "Consistency from individual" for transformation.

#### Process models used with justification SDLC model

##### Software Development Life Cycle (SDLC)

The Software Development Lifecycle (SDLC) for little to medium database application improvement efforts. This adventure uses iterative headway lifecycle, where portions of the application are made through a movement of tight cycle. The primary emphasis center around fundamental usefulness, with resulting cycles adding new usefulness to the past work and additionally revising mistakes distinguished for the parts underway.



### Roles and Responsibilities of PER AND PDR

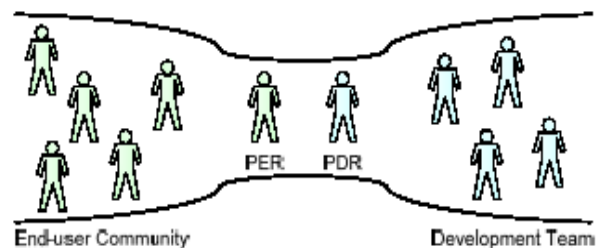
The iterative lifecycle shows two essential employments that exhibition together to unquestionably give adventure issues and thoughts between the end-customer arrange and the improvement gathering.

#### Primary End-user Representative (PER)

The PER is a person who goes about as the basic reason for contact and head approver for the end-customer organize. The PER is in like manner at risk for ensuring that reasonable subject masters lead end-customer reviews in a helpful manner.

#### PDR-PER Relationship

The PER and PDR are the cerebrum trust for the headway effort. The PER has the secret sauce and space data essential to understand the issues related with the



business strategies to the maintained by the application and has a close by working relationship with various people from the end-customer organize. The PDR has comparative central focuses as for the application improvement process and various people from the headway bunch together, they go about as the obsession centers for data about the application to be made.

The objective of this philosophy is to make the comfortable relationship that is typical for an item adventure with one architect and one end-customer essentially, this procedure the "pair programming" thought from Agile approaches and extends it to the end-customer organize. While it is difficult to make comfortable associations between the varying people from an end-customer arrange and an item improvement gathering, it is significantly increasingly direct to make a comfortable association between the lead delegates for each social affair.

### 4. Design Principles & Methodology

#### Object Oriented Design and Analysis

Right when Object bearing is used in assessment similarly as plan, the cutoff among OOA and OOD is

darkened. This is particularly legitimate in methodologies that combine assessment and structure. One clarification behind this darkening is the similarity of basic forms (i.e., objects and classes) that are used in OOA and OOD. Through there is no understanding about what parts of the thing arranged improvement process has a spot with examination and what parts to design, there is some expansive comprehension about the spaces of the two activities.

```

110 x_2 x_3 x_4 x_5 x_6 x_7 x_8 x_9 x_10 x_11 x_12 x_13 x_14 x_15 x_16 x_17 x_18 x_19 x_20 x_21 x_22 x_23 x_24 x_25 x_26 x_27 x_28 x_29 x_30 x_31 x_32 x_33 x_34 x_35 x_36 x_37 x_38 x_39 x_40 x_41 x_42 x_43 x_44 x_45 x_46 x_47 x_48 x_49 x_50 x_51 x_52 x_53 x_54 x_55 x_56 x_57 x_58 x_59 x_60 x_61 x_62 x_63 x_64 x_65 x_66 x_67 x_68 x_69 x_70 x_71 x_72 x_73 x_74 x_75 x_76 x_77 x_78 x_79 x_80 x_81 x_82 x_83 x_84 x_85 x_86 x_87 x_88 x_89 x_90 x_91 x_92 x_93 x_94 x_95 x_96 x_97 x_98 x_99 x_100 x_101 x_102 x_103 x_104 x_105 x_106 x_107 x_108 x_109 x_110 x_111 x_112 x_113 x_114 x_115 x_116 x_117 x_118 x_119 x_120 x_121 x_122 x_123 x_124 x_125 x_126 x_127 x_128 x_129 x_130 x_131 x_132 x_133 x_134 x_135 x_136 x_137 x_138 x_139 x_140 x_141 x_142 x_143 x_144 x_145 x_146 x_147 x_148 x_149 x_150 x_151 x_152 x_153 x_154 x_155 x_156 x_157 x_158 x_159 x_160 x_161 x_162 x_163 x_164 x_165 x_166 x_167 x_168 x_169 x_170 x_171 x_172 x_173 x_174 x_175 x_176 x_177 x_178 x_179 x_180 x_181 x_182 x_183 x_184 x_185 x_186 x_187 x_188 x_189 x_190 x_191 x_192 x_193 x_194 x_195 x_196 x_197 x_198 x_199 x_200 x_201 x_202 x_203 x_204 x_205 x_206 x_207 x_208 x_209 x_210 x_211 x_212 x_213 x_214 x_215 x_216 x_217 x_218 x_219 x_220 x_221 x_222 x_223 x_224 x_225 x_226 x_227 x_228 x_229 x_230 x_231 x_232 x_233 x_234 x_235 x_236 x_237 x_238 x_239 x_240 x_241 x_242 x_243 x_244 x_245 x_246 x_247 x_248 x_249 x_250 x_251 x_252 x_253 x_254 x_255 x_256 x_257 x_258 x_259 x_260 x_261 x_262 x_263 x_264 x_265 x_266 x_267 x_268 x_269 x_270 x_271 x_272 x_273 x_274 x_275 x_276 x_277 x_278 x_279 x_280 x_281 x_282 x_283 x_284 x_285 x_286 x_287 x_288 x_289 x_290 x_291 x_292 x_293 x_294 x_295 x_296 x_297 x_298 x_299 x_300 x_301 x_302 x_303 x_304 x_305 x_306 x_307 x_308 x_309 x_310 x_311 x_312 x_313 x_314 x_315 x_316 x_317 x_318 x_319 x_320 x_321 x_322 x_323 x_324 x_325 x_326 x_327 x_328 x_329 x_330 x_331 x_332 x_333 x_334 x_335 x_336 x_337 x_338 x_339 x_340 x_341 x_342 x_343 x_344 x_345 x_346 x_347 x_348 x_349 x_350 x_351 x_352 x_353 x_354 x_355 x_356 x_357 x_358 x_359 x_360 x_361 x_362 x_363 x_364 x_365 x_366 x_367 x_368 x_369 x_370 x_371 x_372 x_373 x_374 x_375 x_376 x_377 x_378 x_379 x_380 x_381 x_382 x_383 x_384 x_385 x_386 x_387 x_388 x_389 x_390 x_391 x_392 x_393 x_394 x_395 x_396 x_397 x_398 x_399 x_400 x_401 x_402 x_403 x_404 x_405 x_406 x_407 x_408 x_409 x_410 x_411 x_412 x_413 x_414 x_415 x_416 x_417 x_418 x_419 x_420 x_421 x_422 x_423 x_424 x_425 x_426 x_427 x_428 x_429 x_430 x_431 x_432 x_433 x_434 x_435 x_436 x_437 x_438 x_439 x_440 x_441 x_442 x_443 x_444 x_445 x_446 x_447 x_448 x_449 x_450 x_451 x_452 x_453 x_454 x_455 x_456 x_457 x_458 x_459 x_460 x_461 x_462 x_463 x_464 x_465 x_466 x_467 x_468 x_469 x_470 x_471 x_472 x_473 x_474 x_475 x_476 x_477 x_478 x_479 x_480 x_481 x_482 x_483 x_484 x_485 x_486 x_487 x_488 x_489 x_490 x_491 x_492 x_493 x_494 x_495 x_496 x_497 x_498 x_499 x_500 x_501 x_502 x_503 x_504 x_505 x_506 x_507 x_508 x_509 x_510 x_511 x_512 x_513 x_514 x_515 x_516 x_517 x_518 x_519 x_520 x_521 x_522 x_523 x_524 x_525 x_526 x_527 x_528 x_529 x_530 x_531 x_532 x_533 x_534 x_535 x_536 x_537 x_538 x_539 x_540 x_541 x_542 x_543 x_544 x_545 x_546 x_547 x_548 x_549 x_550 x_551 x_552 x_553 x_554 x_555 x_556 x_557 x_558 x_559 x_560 x_561 x_562 x_563 x_564 x_565 x_566 x_567 x_568 x_569 x_570 x_571 x_572 x_573 x_574 x_575 x_576 x_577 x_578 x_579 x_580 x_581 x_582 x_583 x_584 x_585 x_586 x_587 x_588 x_589 x_590 x_591 x_592 x_593 x_594 x_595 x_596 x_597 x_598 x_599 x_600 x_601 x_602 x_603 x_604 x_605 x_606 x_607 x_608 x_609 x_610 x_611 x_612 x_613 x_614 x_615 x_616 x_617 x_618 x_619 x_620 x_621 x_622 x_623 x_624 x_625 x_626 x_627 x_628 x_629 x_630 x_631 x_632 x_633 x_634 x_635 x_636 x_637 x_638 x_639 x_640 x_641 x_642 x_643 x_644 x_645 x_646 x_647 x_648 x_649 x_650 x_651 x_652 x_653 x_654 x_655 x_656 x_657 x_658 x_659 x_660 x_661 x_662 x_663 x_664 x_665 x_666 x_667 x_668 x_669 x_670 x_671 x_672 x_673 x_674 x_675 x_676 x_677 x_678 x_679 x_680 x_681 x_682 x_683 x_684 x_685 x_686 x_687 x_688 x_689 x_690 x_691 x_692 x_693 x_694 x_695 x_696 x_697 x_698 x_699 x_700 x_701 x_702 x_703 x_704 x_705 x_706 x_707 x_708 x_709 x_710 x_711 x_712 x_713 x_714 x_715 x_716 x_717 x_718 x_719 x_720 x_721 x_722 x_723 x_724 x_725 x_726 x_727 x_728 x_729 x_730 x_731 x_732 x_733 x_734 x_735 x_736 x_737 x_738 x_739 x_740 x_741 x_742 x_743 x_744 x_745 x_746 x_747 x_748 x_749 x_750 x_751 x_752 x_753 x_754 x_755 x_756 x_757 x_758 x_759 x_760 x_761 x_762 x_763 x_764 x_765 x_766 x_767 x_768 x_769 x_770 x_771 x_772 x_773 x_774 x_775 x_776 x_777 x_778 x_779 x_780 x_781 x_782 x_783 x_784 x_785 x_786 x_787 x_788 x_789 x_790 x_791 x_792 x_793 x_794 x_795 x_796 x_797 x_798 x_799 x_800 x_801 x_802 x_803 x_804 x_805 x_806 x_807 x_808 x_809 x_810 x_811 x_812 x_813 x_814 x_815 x_816 x_817 x_818 x_819 x_820 x_821 x_822 x_823 x_824 x_825 x_826 x_827 x_828 x_829 x_830 x_831 x_832 x_833 x_834 x_835 x_836 x_837 x_838 x_839 x_840 x
```

The key differentiation among OOA and OOD is that the past models the issue space, provoking an understanding and detail of the issue, while the last models the response for the issue. That is, assessment deals with the issue zone, while design deals with the plan space. Regardless, in OOAD subsumed in the course of action territory depiction. That is, the course of action space depiction, made by OOD, generally contains an incredible piece of the depiction made by OOA. The segregating line is matter of insight, and different people have different points of view on it. The nonattendance of clear division among assessment and design can in like manner be seen as one of the strong motivations behind the article arranged approach the change from examination to setup is "steady". This is furthermore the standard explanation OOAD techniques where examination and plans are both performed.

The standard difference among OOA and OOD, as a result of the different spaces of illustrating, is in the kind of things that leave the assessment and arrangement process.

### Extreme Learning machine

```
20]: srhl_tanh = MLPRandomLayer(n_hidden=10, activation_func='sigmoid')  
elm = GenELMClassifier(hidden_layer=srhl_tanh)  
elm.fit(X_train,y_train)  
pred_elm_model = elm.predict(X_test)  
elm_cm = confusion_matrix(y_test,pred_elm_model)  
accuracy_score(y_test,pred_elm_model)
```

```
20]: 0.7757089345794392
```

```
enter url  
www.google.com  
module 'whois' has no attribute 'whois'  
[[0, -1, 0, -1, 0, -1, -1, 1, 0, 0, 0, -1, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0]]
```

## 5. Conclusion

Right now, phishing ambush features were depicted and we suggested a depiction model to plan the phishing attacks. This technique includes locale extraction features and segment on requests. We have certainly represented concepts of phishing feature extraction in the extraction of the part and these gauges have been used to obtain features. SVM, NB and ELM were used for representation of these bits. Six various starting limits were used in the ELM, and most important precision score was obtained in ELM.

## References

- [1] G. Canbek and Ü. SaðÖro÷lu, "A Review on Information, Information Security and Security Processes," *Politek. Derg.*, vol. 9, no. 3, pp. 165–174, 2006.
- [2] L. McCluskey, F. Thabtah, and R. M. Mohammad, "Intelligent rule- based phishing websites classification," *IET Inf. Secur.*, vol. 8, no. 3, pp. 153–160, 2014.
- [3] R. M. Mohammad, F. Thabtah, and L. McCluskey, "Predicting phishing websites based on self-structuring neural network," *Neural Comput. Appl.*, vol. 25, no. 2, pp. 443–458, 2014.
- [4] R. M. Mohammad, F. Thabtah, and L. McCluskey, "An assessment of features related to phishing websites using an automated technique," *Internet Technol. ...*, pp. 492–497, 2012.
- [5] W. Hadi, F. Aburub, and S. Alhawari, "A new fast associative classification algorithm for detecting phishing websites," *Appl. Soft Comput. J.*, vol. 48, pp. 729–734, 2016.
- [6] N. Abdelhamid, "Multi-label rules for phishing classification," *Appl. Comput. Informatics*, vol. 11, no. 1, pp. 29–46, 2015.
- [7] N. Sanglerdsinlapachai and A. Rungsawang, "Using domain top-page similarity feature in

machine learning-based web phishing detection,” in 3rd International Conference on Knowledge Discovery and Data Mining, WKDD 2010, 2010, pp. 187–190.

[8] W. D. Yu, S. Nargundkar, and N. Tiruthani, “A phishing vulnerability analysis of web based systems,” IEEE Symp. Comput. Commun. (ISCC 2008), pp. 326–331, 2008.

[9] P. Ying and D. Xuhua, “Anomaly based web phishing page detection,” in Proceedings - Annual Computer Security Applications Conference, ACSAC, 2006, pp. 381–390.

[10] M. Moghimi and A. Y. Varjani, “New rule-based phishing detection method,” Expert Syst. Appl., vol. 53, pp. 231–242, 2016.

[11] DATASET: Lichman, M. (2013). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California, School of Information and Computer Science

[12] G.-B. Huang et al., “Extreme learning machine: Theory and applications,” Neurocomputing, vol. 70, no. 1–3, pp. 489–501, 2006.

[13] C. S. Guang-bin Huang, Qin-yu Zhu, “Extreme learning machine: A new learning scheme of feedforward neural networks,” Neurocomputing, vol. 70, pp. 489–501, 2006.

[14] T. S. Guzella and W. M. Caminhas, “A review of machine learning approaches to Spam filtering,” Expert Systems with Applications, vol. 36, no. 7. pp. 10206–10222, 2009.

[15] Ö.F..Ertu÷rul,AúÖrÖÖ÷renmeMakineleriilebiyo lojiksinyalleringizlikaynaklarÖnaayrÖútÖrÖlmasÖ. D.Ü. MühendislikDergisiCilt: 7, 1, 3-9-2016

[16] M. E. Tagluk, M. S. Mamiú, M. Arkan, and Ö. F. Ertugrul, “AúiriÖgrenmeMakineleriileEnerjiletimHatları Ariza Tipi veYerininTespiti,” in 2015 23rd Signal Processing and Communications Applications Conference, SIU 2015 - Proceedings, 2015, pp. 1090– 1093.

[17] Ö. Faruk Ertu÷rul and Y. Kaya, “A detailed analysis on extreme learning machine and novel approaches based on ELM,” Am. J. Comput. Sci. Eng., vol. 1, no. 5, pp. 43–50, 2014.

[18] Ö. F. Ertugrul, “Forecasting electricity load by a novel recurrent extreme learning machines approach,” Int. J. Electr. Power Energy Syst., vol. 78, pp. 429–435, 2016.

[19] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, “Extreme learning machine: Theory and applications,” Neurocomputing, vol. 70, no. 1, pp. 489–501,2006.