

برنامه ریزی فرآیند تعمیر و نگهداری در محیط تولیدی که از وسایل نقلیه هدایت شونده خودکار (*Automatic guided vehicle*) برای انتقال مواد اولیه و محصول نیم ساخته و محصول نهایی بین ایستگاه های کاری و انبار محصول نهایی استفاده می گردد. این محیط تولیدی شامل ۷ ایستگاه کاری و ۳ *AGV* می باشد.

همچنین موقعیتی تحت عنوان سایت نت باید در نظر گرفته شود که در واقع این محل می تواند با محل دپو *AGV* ها (محلی که *AGV* ها شارژ می شوند) یکی باشد و یا به صورت جداگانه باشد و همچنین سایت تعمیر و نگهداری *AGV* ها می تواند با سایت نت دستگاه های تولیدی به صورت مشترک یا جداگانه در نظر گرفته شود که شرایط ذیل برای این منظور بایستی در نظر گرفته شود:

A. شامل موقعیت سایت نت + دپو جداگانه برای *AGV* ها که در کنار خط تولید مستقر گردیده است (زمان انتقال از سایت به ایستگاه های کاری / انبار کمتر می باشد).

B. سایت نت مشترک با دیگر دستگاه ها و ماشین آلات تولید و قرار گرفتن دپو در همان محل که در محلی جدا از خط تولید می باشد. (زمان انتقال بین سایت و ایستگاه های کاری / انبار بیشتر می باشد).

C. و یا سایت نت مشترک با دیگر دستگاه های تولیدی ولی محل دپو در کنار خط تولید (زمان انتقال بین دپو و ایستگاه های کاری / انبار مانند حالت A می باشد. و زمان انتقال بین سایت نت و ایستگاه های کاری / انبار مانند حالت B می باشد).

Condition	AGV Depot & Maintenance AGV	AGV Depot	PM Maintenance	AGV & PM Maintenance & Depot
A	✓	-	✓	-
B	-	-	-	✓
C	-	✓	-	✓

مدت زمان انتقال بین ایستگاه های کاری و انبار و سایت نت و دپو از قبل تعیین شده است. در صورتیکه *AGV* ها در حالت *Abnormal* قرار گیرند این مدت زمان انتقال با ضریب ثابتی نسبت به حالت عملکرد افزایش می یابد زیرا فرض می گردد که *AGV* ها در این حالت برای حفظ ایمنی با سرعت کمتری حرکت می نمایند. که این کاهش سرعت باعث افزایش ۴۰ درصدی زمان انتقال بین دو واحد می گردد. (انتقال می تواند از انبار به ایستگاه کاری یا از ایستگاه کاری به ایستگاه کاری دیگر و از ایستگاه کاری به انبار و یا از دپو به هر ایستگاه کاری و انبار باشد).

Workstation and Stock distance	Goods Storage	W1	W2	W3	W4	W5	W6	W7	Product Storage	Depot
Goods Storage		0.3	0.4	0.3	0.8	0.1	0.7	0.9	1	0.1
W1	0.3		0.5	0.4	0.1	0.2	0.7	0.8	0.7	0.5
W2	0.4	0.5		0.4	0.4	0.5	1	1	0.9	0.3
W3	0.3	0.4	0.4		0.7	0.4	0.1	1	0.4	0.4

W4	0.8	0.1	0.4	0.7	0.3	0.7	0.2	0.8	0.3
W5	0.1	0.2	0.5	0.4	0.3	0.9	0.5	0.9	0.5
W6	0.7	0.7	1	0.1	0.7	0.9	0.4	0.8	0.1
W7	0.9	0.8	1	1	0.2	0.5	0.4	0.3	0.3
Product Storage	1	0.7	0.9	0.4	0.8	0.9	0.8	0.3	0.4
Depot	0.1	0.5	0.3	0.4	0.3	0.5	0.1	0.3	

تعداد سفارش های موجود شده بر روی بافر در هر ساعت به صورت تصادفی و براساس توزیع احتمال یکنواخت بوجود می آید.

سیاست ها نت در مجموع شامل ۳ نوع سیاست اصلاحی، پیشگیرانه و پیشگویانه می باشند. (همانطور که در بالا توضیح داده شد باید در هنگام اجرا برنامه پرسیده شود که موقعیت سایت نت در کدام موقعیت بایستی قرار گیرد و همچنین کدام سیاست های نت (اصلاحی، پیشگیرانه، پیشگویانه) بایستی در نظر گرفته شود).

میانگین زمان کارکرد تا خرابی (MTTF) برای ماشین آلات تولیدی از توزیع نرمال با میانگین ۴۰۰ ساعت و انحراف معیار ۱۲۰ ساعت پیروی میکند.

همچنین مدت زمان تعمیر ماشین آلات تولیدی در حالت اصلاحی، پیشگیرانه و پیشگویانه برای هر ماشین از توزیع یکنواخت  $U(1,15)$  ساعت پیروی میکند و نوع تعمیر ماشین آلات تولیدی مهم نبوده و فقط مدت زمان نیاز آن ها به تکنسین های نت مورد هدف است.

AGV ها از ۳ قسمت کلی مطابق مقاله یان و همکاران ۲۰۱۷ تشکیل شده اند ۱. قسمت های مکانیکال (موتور، ترمز، سیستم فرمان و هدایت AGV و اتصالات فیزیکی ۲. سیستم (شامل قسمت های سیستم مسیریابی لیزری، سیستم ایمنی و سیستم نرم افزار کنترل AGV) ۳. باتری تشکیل شده اند.

هدف مقاله این هست که بهترین زمان انجام فرآیند های نت پیشگویانه را با توجه به حجم کار و مدت زمان باقی مانده تا خرابی اصلاحی و بیکاری تکنسین ها را با استفاده از یادگیری ماشین تقویتی بدست آورد و همچنین بهترین زمان انجام فرآیند های نت پیشگیرانه را با توجه به حجم بافر ها و بیکاری تکنسین های نت در نظر گیرد.

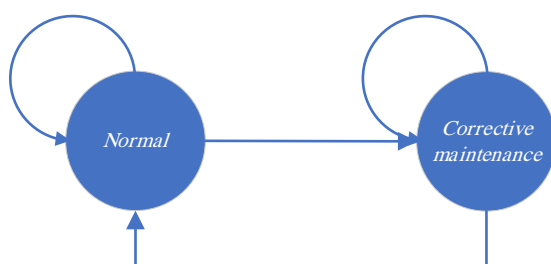
برای این منظور بایستی با دو الگوریتم یادگیری ماشین تقویتی ۳ سیاست تعمیرات و نگهداری و ۳ شرایط مختلف قرار گیری سایت تعمیرات و نگهداری و دپو در نظر گرفته شود و پیاده سازی گردد بنابراین مطابق جدول ذیل بایستی در نظر گرفته شوند:

<i>Models</i>	<i>Maintenance Policy</i>	<i>Condition Location</i>
<i>1st Model</i>	<i>CM</i>	<i>A</i>
<i>2nd Model</i>	<i>CM</i>	<i>B</i>
<i>3rd Model</i>	<i>CM</i>	<i>C</i>
<i>4th Model</i>	<i>CM, PM</i>	<i>A</i>
<i>5th Model</i>	<i>CM, PM</i>	<i>B</i>
<i>6th Model</i>	<i>CM, PM</i>	<i>C</i>
<i>7th Model</i>	<i>CM, PM, PdM</i>	<i>A</i>
<i>8th Model</i>	<i>CM, PM, PdM</i>	<i>B</i>
<i>9th Model</i>	<i>CM, PM, PdM</i>	<i>C</i>
<i>PdM = Predictive Maintenance, PM= Preventive Maintenance, CM= Corrective Maintenance</i>		

### حالت اول: در نظر گرفتن تعمیرات و نگهداری اصلاحی:

در این حالت تنها تعمیرات و نگهداری اصلاحی برای وسایل نقلیه شونده خودکار پیاده سازی می شوند، بنابراین حالت ها شامل دو حالت می باشند که سالم و خراب بودن دستگاه را نشان می دهد احتمال انتقال از وضعیت سالم به خراب با توجه به احتمال خرابی AGV ها که از یک توزیع وایبل با پارامتر مقیاس ۱۴۰ و ۱۹۵ ساعت به ترتیب برای قسمت های الکتریکال و مکانیکال با پارامتر شکل متغیر ۰/۵ و ۵ با توجه به زمان پیروی می کند و برای قسمت باتری از یک توزیع نمایی با میانگین ۵۰۰ ساعت پیروی می نماید و همچنین شرایط زمان تخلیه باتری و نیاز به شارژ دستگاه با توجه به میزان حجم بار بطور میانگین از یک توزیع نرمال با میانگین ۳۰ ساعت و واریانس ۲ ساعت پیروی میکند یعنی مدت زمانی که طول می کشد میزان شارژ باتری از ۱۰۰٪ به ۰٪ برسد بطور متوسط ۳۰ ساعت می باشد و مدت زمان شارژ شدن کامل باتری ۱۰ ساعت می باشد همچنین فرض می شود بعد از هر تعمیر دستگاه به حالت اولیه و کاملاً سالم بر می گردد. مدت زمان انجام تعمیر و نگهداری اصلاحی از توزیع نرمال  $N(22,3)$  برای هر قسمت پیروی می نماید.  $i$  در فرمول پایین واحد زمانی بوده که نشان دهنده وضعیت دستگاه در هر ساعت می باشد. در این حالت با توجه به حالت Run-to-failure بودن فرآیند تعمیر و نگهداری، هیچ گونه فرآیندی جهت برنامه ریزی وجود نداشته و تنها بایستی شبیه سازی صورت پذیرد.

$$S_{t=i} = \text{Normal or Corrective Maintenance}$$



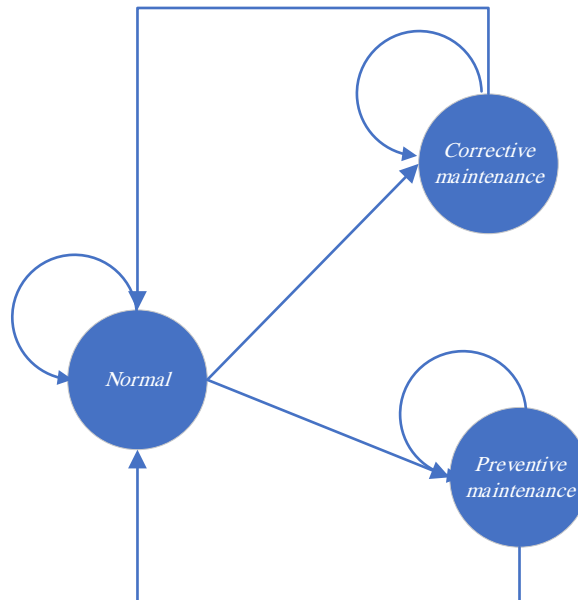
در این حالت میزان هزینه تعمیرات و نگهداری شامل هزینه قطعات یدکی مصرفی که به صورت تصادفی عددی بین ۰ تا ۴۰۰۰ واحد پولی می باشد بوده است (یعنی هر بار که دستگاه در حالت CM قرار می گیرد هزینه ای به عنوان هزینه قطعات یدکی به سیستم تحمیل می گردد) و هزینه مربوط به دستمزد تکنسین تعمیر و نگهداری که به ازای هر ساعت ۱۰۰ واحد پولی بوده و از ضرب هزینه هر ساعت دستمزد در مدت زمانی که دستگاه ها در حالت CM می باشند بدست می آید.

### حالت دوم: در نظر گرفتن تعمیرات و نگهداری اصلاحی به همراه تعمیرات و نگهداری پیشگیرانه

در این حالت تنها تعمیرات و نگهداری اصلاحی مانند حالت اول در نظر گرفته می شود و برای تعمیر و نگهداری پیشگیرانه هر ۱۶۵ ساعت یکبار بدون در نظر گرفتن وضعیت دستگاه سالم یا خراب صورت می پذیرد که مدت زمان آن ۱ ساعت خواهد بود که با توجه به کیفیت تعمیر و نگهداری پیشگیرانه صورت پذیرفته دستگاه به یکی از حالت ها در شرایط قبل بر می گردد. که در این صورت با توجه به مرحله زمانی که در آن قرار دارد باید به عنوان مثال مرحله زمانی  $m$  باشد در این صورت با توجه به کیفیت تعمیر  $r$  که یک عدد تصادفی بین ۰ تا  $m-1$  می باشد دستگاه به  $m-r$  مرحله قبل بر خواهد گشت.

$$S_{t=i} = \text{Normal or CM or PM}$$

در اینجا عامل باید دو عمل *Continue* یا *PMRepair* را انتخاب نماید که بایستی آن را با توجه به حجم بافر ها و مدت زمان بیکاری تکنسین ها بعد از ۱۶۵ ساعت کارکرد انتخاب نماید.



پاداش در این حالت به ۳ مورد حجم بافر و بیکاری تکنسین ها و زمان انجام به موقع تعمیرات و نگهداری پیشگیرانه بستگی دارد و عامل بایستی با توجه به این ۳ مورد فرآیند تعمیرات و نگهداری پیشگیرانه را انجام دهد. همچنین چنانچه به دلیل عدم به موقع

انجام دادن برنامه ریزی تعمیرات و نگهداری پیشگیرانه دستگاه خراب شود و برای تعمیرات و نگهداری اصلاحی فرستاده شود، عامل یک امتیاز منفی بزرگ دریافت می کند. بنابراین با توجه به موارد ذکر شده پاداش با توجه به فرمول های ذیل بدست می آید:

- حجم بافر:

$BV$  = حجم بافر

$M$  = ماکزیمم حجم بافر

$$R1_{PM} = 1 - \left(\frac{BV}{M}\right)$$

- بیکاری تکنسین ها:

$$R2_{PM} = \begin{cases} 1 & \text{اگر یک نفر یا هر دو نفر از تکنسین ها واحد تعمیرات و نگهداری بیکار باشند.} \\ 0 & \text{در غیر این صورت} \end{cases}$$

- زمان اجرای فرآیند تعمیرات و نگهداری پیشگیرانه:

$t_{RPM}$  زمان واقعی که فرآیند تعمیرات و نگهداری پیشگیرانه صورت پذیرفته است =

$t_{PPM}$  زمان برنامه ریزی شده که فرآیند تعمیرات و نگهداری پیشگیرانه باید در آن زمان صورت می پذیرفته است =

$$R3_{PM} = \frac{1}{t_{RPM} - t_{PPM}}$$

- زمان انجام فرآیند تعمیرات و نگهداری اصلاحی: چنانچه فرآیند تعمیرات و نگهداری اصلاحی بعد از  $t_{PPM}$  اتفاق بیافتد

در حالیکه  $t_{PPM}$  هنوز انجام نشده باشد عامل به اندازه ۱۰- جریمه می شود ولی چنانچه قبل از  $t_{PPM}$  اتفاق بیافتد یا

بعد از  $t_{PPM}$  اتفاق بیافتد در حالیکه  $t_{PPM}$  انجام نشده باشد عامل میزان پاداش صفر را دریافت می نماید.

$$R4_{PM} = \begin{cases} -10 & \text{After } t_{PPM} \text{ and } t_{PPM} \text{ did not perform} \\ 0 & \text{در غیر این صورت} \end{cases}$$

در این حالت میزان هزینه تعمیرات و نگهداری اصلاحی مانند حالت قبل بوده و هزینه تعمیرات و نگهداری پیشگیرانه شامل هزینه قطعات یدکی مصرفی که به ازای هر بار اجرای تعمیرات و نگهداری پیشگیرانه عددی تصادفی بین ۰ تا ۵۰۰ واحد پولی انتخاب می گردد و میزان دستمزد مانند حالت قبل از حاصلضرب ۱۰۰ واحد پولی به عنوان میزان دستمزد تکنسین ها به ازای یک ساعت در مدت زمانی که دستگاه ها تحت تعمیر و نگهداری پیشگیرانه قرار گرفته اند محاسبه می گردد.

#### حالت سوم: در نظر گرفتن سیاست های تعمیرات و نگهداری اصلاحی به همراه تعمیرات و نگهداری پیشگیرانه و نگهداری پیشگویانه

برای این منظور احتمال اعلام خطا در سنسور های قسمت های مختلف الکتریکال، مکانیکال و باطری از یک توزیع وایبل پیروی میکند که بعد از خارج شدن عدد سنسور از محدوده تعریف شده دستگاه به حالت عملکرد غیرعادی (Abnormal) یا اضطراری رفته و طبق توزیع نرمال مدت زمان تا خرابی دستگاه مشخص می گردد (که میزان خرابی دستگاه فزاینده می باشد) و در نتیجه با توجه به آن باید دستگاه در اولین زمان ممکن که حجم کار روی بافر آن کم و تکنسین های نت بیکار می باشند از چرخه عملیاتی خارج و تعمیرات پیشگویانه روی آن صورت پذیرد. به همین منظور میانگین زمان کارکرد تا نقص (MTTD) برای AGV ها با توجه به جدید بودن آن ها از یک توزیع وایبل با پارامتر مقیاس ۱۰۰ و ۱۵۰ ساعت به ترتیب برای قسمت های الکتریکال و مکانیکال و پارامتر شکل متغیر ۵/۰ و ۵ با توجه به زمان پیروی می کند و زمان تخلیه باتری و نیاز به شارژ دستگاه با توجه به میزان حجم بار بطور میانگین از یک توزیع نرمال با میانگین ۲۵ ساعت و واریانس ۲ ساعت پیروی میکند یعنی مدت زمانی که دستگاه طول می کشد که میزان شارژ باتری از ۱۰۰٪ به ۲۰٪ برسد بطور متوسط ۲۵ ساعت می باشد. پس از خارج شدن سیستم از حالت Normal و رفتن به حالت Abnormal مدت زمان عملکرد سیستم تا خرابی در این حالت برای قسمت های الکتریکال و مکانیکال برای هر سنسور متفاوت بوده و از توزیع نرمال با میانگین و واریانس های مربوط به خود پیروی می کند که در جدول ذیل قابل مشاهده می باشد.

<i>Part</i>	<i>Sub-Part</i>	<i>Sensors</i>	<i>Distribution in abnormal states</i>	<i>Predictive Repair time/Charge (Normal Distribution)</i>	<i>Predictive Replace time (Normal Distribution)/ Charge</i>	<i>Preventive Repair time</i>	<i>Corrective Repair time</i>
<i>Mechanical</i>	<i>Motors, Steering Control and Attachments</i>	<i>Energy Consumption</i>	<i>N(14,5)</i>	<i>(15;7)</i>	<i>(4;1)</i>	<i>1</i>	<i>(23;1)</i>
		<i>Vibration</i>	<i>N(32,3)</i>	<i>(8;3)</i>	<i>(3;1)</i>		<i>(17;2)</i>
		<i>Noise</i>	<i>N(25.2)</i>	<i>(14;7)</i>	<i>(2;1)</i>		<i>(19;1)</i>
		<i>Temperature</i>	<i>N(15,4)</i>	<i>(13;4)</i>	<i>(5;1)</i>		<i>(25;1)</i>
	<i>Brake</i>	<i>Electrical brake</i>	<i>N(10,3)</i>	<i>(10;2)</i>	<i>(5;1)</i>		<i>(24;2)</i>
		<i>Mechanical brake</i>	<i>N(10,3)</i>	<i>(10;3)</i>	<i>(5;1)</i>		<i>(24;2)</i>
<i>Systems</i>	<i>Navigation &amp; System control &amp; Safety</i>	<i>Laser sensors/Emitters</i>	<i>N(5,2)</i>	<i>(15;3)</i>	<i>(5;1)</i>		<i>(21;2)</i>
		<i>Visual</i>	<i>N(20,5)</i>	<i>(10;2)</i>	<i>(4;1)</i>		<i>(21;1)</i>
<i>Battery</i>	<i>Percentage of charge</i>	<i>Percentage of charge</i>	<i>N(6,1)</i>	<i>Due to have charge (8;1)</i> -			<i>10</i>

کیفیت تعمیر در سیاست پیشگیرانه مانند حالت دوم می باشد ولی در سیاست پیشگیرانه بازه انتخابی  $r$  از ۱ تا  $m-1$  بوده است یعنی با توجه به کیفیت تعمیر دستگاه به یکی از مراحل بین مرحله قبل از تعمیر پیشگیرانه و یا به حالت بعد از مرحله اولیه خواهد برگشت. همچنین فرآیند تعمیر و نگهداری پیشگیرانه در هر ۱۲۰ ساعت باید با توجه حجم بافر و بیکاری تکنسین های تعمیر و نگهداری صورت پذیرد.

بازه ایده آل برای هر سنسور در جدول ذیل نمایش داده شده است:

<i>Part</i>	<i>Sub-Part</i>	<i>Sensors</i>	<i>Safe Domain</i>	<i>Critical Domain</i>	<i>Description</i>
<i>Mechanical</i>	<i>Motors, Steering Control and Attachments</i>	<i>Energy Consumption</i>	<i>0 to 1000 Ws</i>	<i>above 1000ws</i>	<i>27.5 kg with speed 1.5 m/s, speed decrease to 0.6 m/s in abnormal states</i>
		<i>Vibration</i>	<i>1.8 mm/s to 2.8 mm/s</i>	<i>2.9 mm/s to failure</i>	-
		<i>Noise</i>	<i>20 to 30 dB</i>	<i>above 30 dB</i>	-
		<i>Temperature</i>	<i>20 to 40</i>	<i>40 to 60</i>	-
	<i>Brake</i>	<i>Electrical brake</i>	<i>1.95 to 2.05</i>	<i>0.95 to 1.05 meter</i>	-
		<i>Mechanical brake</i>	<i>0.95 to 1.05</i>	<i>0 to 0.94</i>	-
<i>Systems</i>	<i>Navigation &amp; System control &amp; Safety</i>	<i>Laser sensors/Emitters</i>	<i>ok or not ok</i>	<i>not ok to failure</i>	-
		<i>Visual</i>	<i>ok or not ok</i>	<i>not ok to failure</i>	-

Battery	Percentage of charge	Percentage of charge	above 20%	Between 0 to 20	-
---------	----------------------	----------------------	-----------	-----------------	---

حالت ها بیانگر عدد سنسورها می باشند و به صورت یک بردار با ۶ درایه بیان می گردد:

$$S_{t=i} = [EC, Vb, N, T, EB, MB, LS, V]$$

$EC=Electrical Consumption$

$Vb=Vibration$

$N=Noise$

$T=Temperature$

$EB=Electrical Brake$

$MB=Mechanical Brake$

$LS=Laser Sensors/Emitters$

$V=Visual$

$i$  نشان دهنده مرحله زمانی که برابر  $N$  مدت زمان شبیه سازی سیستم می باشد و درایه ها نشان دهنده عدد سنسور می باشند که با روش های توضیح داده شده ایجاد خواهند شد و با توجه به عدد بالا تولید خواهند گردید.

عامل ۴ عمل  $action$  در همه حالت ها  $states$  می تواند انجام دهد ۱. تعمیر پیشگویانه ( $PdMRepair$ )

۲. تعویض ( $PdMReplace$ ) ۳. حفظ یا ادامه کار قطعات ( $hold/continue$ ) ۴. تعمیرات پیشگیرانه ( $PMRepair$ )

با توجه به حالت چنانچه دستگاه در عملکرد صحیح باشد و هر یک از عمل های تعمیر و تعویض صورت پذیرد پاداش منفی دریافت می کند و همچنین اگر در حالت  $Abnormal$  باشد و به موقع یکی از عمل های تعمیر یا تعویض را انتخاب نکند و دستگاه به دلیل خرابی متوقف گردد عامل مجددا پاداش منفی دریافت می کند.

برای انتخاب عمل های تعمیر/تعویض درحالت  $Abnormal$  باید این انتخاب با توجه به محدودیت بودجه تعمیر و نگهداری انتخاب شوند. شایان ذکر است چنانچه عمل تعویض باشد دستگاه به حالت اولیه باز می گردد و چنانچه عمل تعمیر باشد دستگاه با توجه به کیفیت تعمیر به یکی از حالت های دستگاه در شرایط نرمال به غیر از حالت اولیه باز می گردد. یعنی آنکه با توجه به حجم بافر و بیکاری تکنسین های نت و عملکرد کاری آن ها فرآیند تعمیر را برنامه ریزی نماید.

فرآیندهای نت برای عمل حفظ دارای هزینه صفر می باشد و در عمل تعمیر و تعویض به ترتیب دارای میانگین هزینه تعمیر بیشتری می باشند که این هزینه برای هریک از توزیع یکنواخت پیروی می کند.

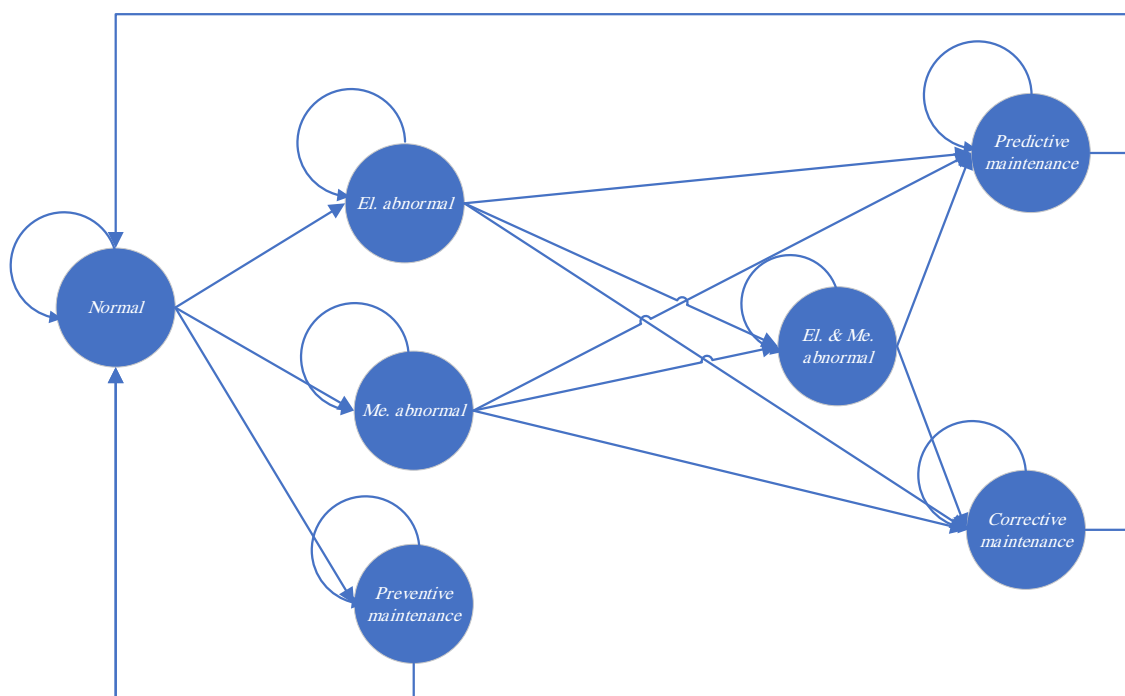
$Agent$  وظیفه برنامه ریزی فرآیند تعمیر و نگهداری را برای این  $AGV$  ها را برعهده دارد و بایستی بهترین زمان انجام فرآیند های نت پیشگویانه را با توجه به حجم کار و مدت زمان باقی مانده تا خرابی اصلاحی و بیکاری تکنسین ها را با استفاده از

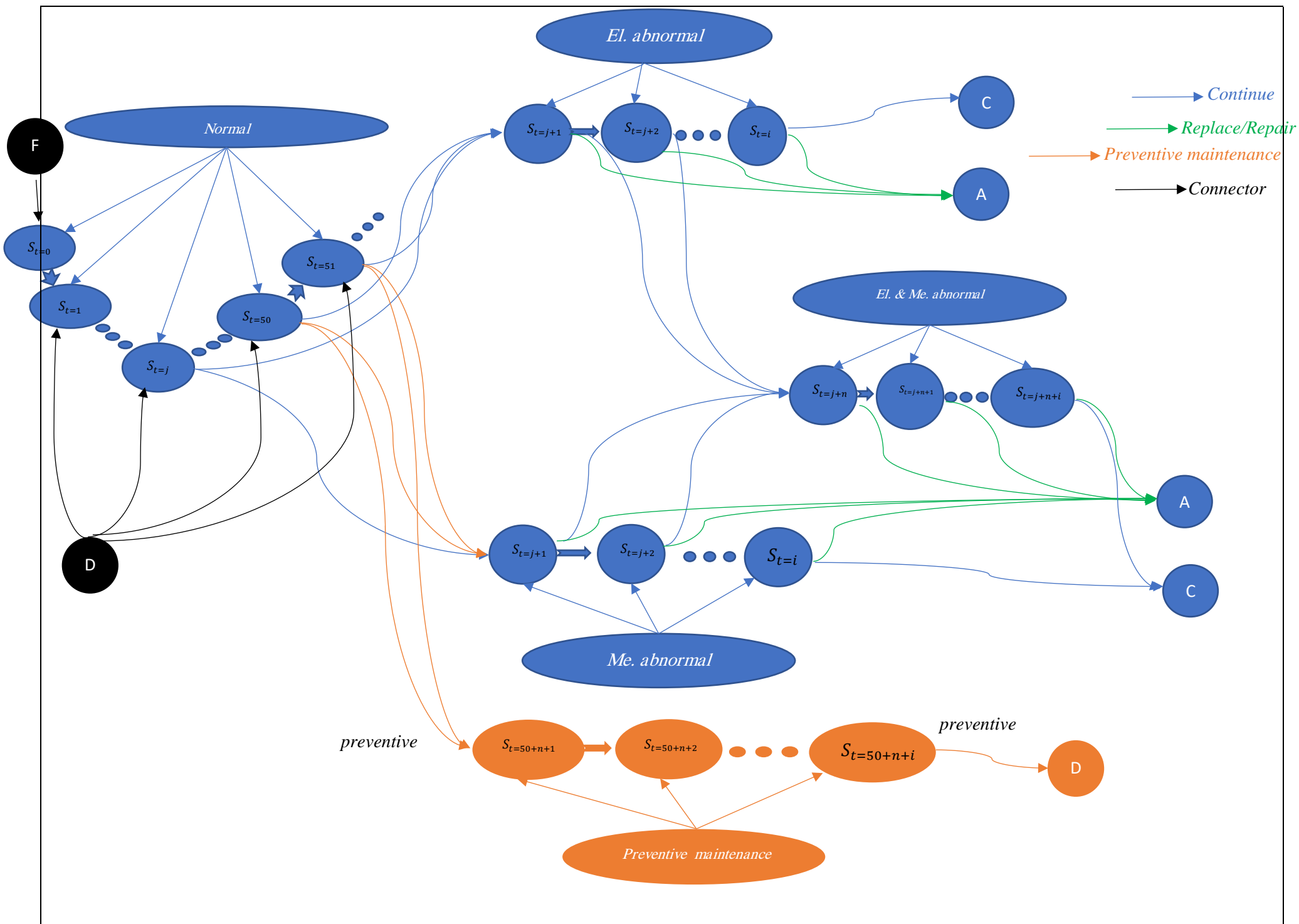


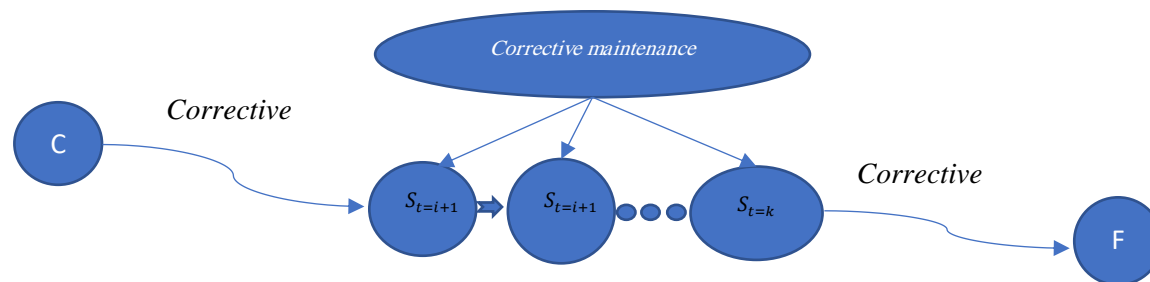
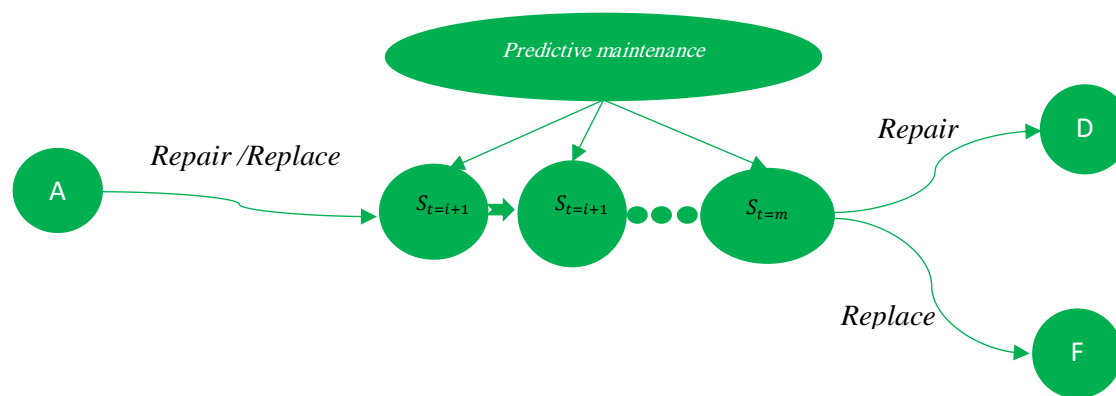
یادگیری ماشین تقویتی بدست آورد و همچنین بهترین زمان انجام فرآیند های نت پیشگیرانه را با توجه به حجم بافر ها و بیکاری تکنسین های نت در نظر گیرد. عامل شرایط ماشین ها را با استفاده از پایش وضعیت (*Conditional monitoring*) در نظر می گیرد که برای پایش وضعیت قسمت الکتریکال از سنسورهایی نظیر مشاهده (*Visual*)، جریان مصرفی (*Current Consumption*) و توان موتور الکتریکی (*Power*) و برای قسمت های مکانیکال از سنسورهایی نظیر ارتعاشی (*Vibration*)، صدا (*noise*) و دما (*temperature*) استفاده کرد که نشان دهنده حالت یا *state* است که *AGV* مربوطه در آن قرار می گیرد. هر حالت به صورت یک بردار شامل درایه های مربوط به اعداد سنسور ها تعریف شده و زمانی یک حالت *Abnormal* در نظر گرفته میشود که یکی از درایه های مربوط به سنسورها از دامنه تعریف شده در جدول ۱ خارج شود.

نحوه محاسبه درایه های داخل بردارها بدین صورت است که با توجه به توزیع وایبل احتمال بقا در زمان مورد نظر بدست آمده و سپس با استفاده از توزیع یکنواخت یک عدد تصادفی بین صفر و یک ایجاد می شود و پس از آن در صورت کمتر مساوی بودن عدد مورد نظر از احتمال بقا عددی تصادفی از داخل بازه تعریف شده به تصادف انتخاب می شود ولی چنانچه عدد تصادفی از احتمال بقا بزرگتر باشد عددی تصادفی از حد بالای بازه تا ماکزیمم حالت انتخاب شده و برای آن سنسور در نظر گرفته می شود.

بطور کلی حالت ها و عمل های در نظر گرفته شده به صورت ذیل می باشند:







پاداش در این حالت به ۲ مورد حجم بافر و بیکاری تکنسین ها و زمان انجام به موقع تعمیرات و نگهداری پیشگیرانه بستگی دارد و عامل بایستی با توجه به این ۲ مورد فرآیند تعمیرات و نگهداری پیشگیرانه را انجام دهد. همچنین چنانچه به دلیل عدم به موقع انجام دادن برنامه ریزی تعمیرات و نگهداری پیشگویانه دستگاه خراب شود و برای تعمیرات و نگهداری اصلاحی فرستاده شود، عامل یک امتیاز منفی بزرگ دریافت می کند. بنابراین با توجه به موارد ذکر شده پاداش با توجه به فرمول های ذیل بدست می آید:

- حجم بافر:

BV= حجم بافر

M= ماکزیمم حجم بافر

$$R1_{pdM} = 1 - \left(\frac{BV}{M}\right)$$

- بیکاری تکنسین ها:

$$R2_{pdM} = \begin{cases} 1 & \text{اگر یک نفر یا هر دو نفر از تکنسین ها واحد تعمیرات و نگهداری بیکار باشند.} \\ 0 & \text{در غیر این صورت} \end{cases}$$

در این حالت میزان هزینه تعمیرات و نگهداری پیشگیرانه مانند حالت قبل بوده و هزینه تعمیرات و نگهداری پیشگویانه برای عمل تعمیر شامل هزینه تنها دستمزد می گردد که همان ۱۰۰ واحد پولی دستمزد در مدت زمانی که دستگاه ها در حالت تعمیر پیشگویانه با عمل تعمیر قرار می گیرند بوده است و برای عمل تعویض در این صورت قطعات یدکی مصرفی که به ازای هر بار اجرای تعمیرات و نگهداری پیشگویانه با عمل تعویض عددی تصادفی بین ۵۰۰ تا ۲۰۰۰ واحد پولی انتخاب می گردد و میزان دستمزد مانند حالت قبل از

حاصلضرب ۱۰۰ واحد پولی به عنوان میزان دستمزد تکنسین ها به ازای یک ساعت در مدت زمانی که دستگاه ها تحت تعمیر و نگهداری پیشگویانه با عمل تعویض قرار گرفته اند محاسبه می گردد و چنانچه دستگاه در حالت تعمیر و نگهداری اصلاحی قرار گیرد، میزان هزینه تعمیرات و نگهداری شامل هزینه قطعات یدکی مصرفی که به صورت تصادفی عددی بین ۱۵۰۰ تا ۴۰۰۰ واحد پولی می باشد بوده است و هزینه مربوط به دستمزد تکنسین تعمیر و نگهداری که به ازای هر ساعت ۱۰۰ واحد پولی بوده و از ضرب هزینه هرساعت دستمزد در مدت زمانی که دستگاه ها در حالت CM می باشند بدست می آید.

در انتها بایستی کارایی برنامه نوشته شده تست شده و سپس بر روی مسئله تعریف شده پیاده سازی شود، همچنین بایستی روش های مختلف یادگیری تقویتی در پیاده سازی گردد(مانند DQN, SARSA, Policy Gradient یا هر روش دیگر پیشنهادی ....) و سپس به بررسی خروجی این روش ها با روش های رایج در مقالات بپردازد.